Improving Mean Forecast with Density Forecast: A Functional Regression Approach*

Yoosoon Chang† Kerry Loaiza-Marín‡ Michael W. McCracken§

Indiana University Indiana University Federal Reserve Bank of St. Louis

Joon Y. Park¶

Indiana University

December 11, 2025

Abstract

We propose a method to improve a given conditional mean forecast with a given density forecast by estimating functional regressions. Our method offers a feasible and rigorous way of aligning conditional mean forecasts with the underlying distributional risks or uncertainties. It relies on functional principal components (FPCs) as an effective way to summarize and exploit the predictive content of density forecasts. We perform an application to forecasting the United States quarterly GDP growth rate. We consider the Survey of Professional Forecasters (SPF)' consensus (average) forecast as a point forecast. Our results show more than 30% improvement in the out-of-sample (OOS) root mean square (forecasting) error for one-year-ahead GDP growth rate by using our functional regression forecasts relative to the SPF's average. The main reason is the predictive content associated with recession and expansion periods which helps decreasing the OOS bias. Our findings suggest that the SPF's consensus (average) forecast tends to be conservative.

Key words: growth-at-risk, real-time data, quantiles, functional regression, nonparametrics, density forecasting.

JEL Codes: C32, C38, C51, C52, C53.

*For helpful comments, we are grateful to Christian Matthes, Robin Braun, Yuan Liao, Matias D. Cattaneo, and seminar participants at the 2025 Hoosier Conference, the Department of Economics of Universidad de Costa Rica, the 2025 Midwest Econometrics Group Conference, and the Banco Central de Costa Rica Research Department for helpful comments. The views expressed here are those of the individual authors and do not necessarily reflect official positions of the Banco Central de Costa Rica, Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

†Bloomington, IN 47405-7104. E-mail: yoosoon@iu.edu

‡Bloomington, IN 47405-7104. E-mail: kloaizam@iu.edu

§St. Louis, MO 63166. Email: michael.w.mccracken@stls.frb.org

¶Bloomington, IN 47405-7104. E-mail: joon@iu.edu

"My approach to this problem [uncertainty over model selection] while on the Federal Reserve Board was relatively simple: (...) simulate a policy on as many of these models as possible, throw out the outlier(s), and average the rest to get a point estimate of a dynamic multiplier path." (Blinder, 1999)

1 Introduction

Forecasts are fundamental inputs for decision making. Governments, central banks, and private sector agents take intertemporal decisions about consumption, savings, investment, taxes, interest rates, among many others. These decisions involve a forward looking perspective which needs the approximation of future values of key variables, being the real Gross Domestic Product (GDP) growth one of the most important. The point estimates of expected values coming from judgmental-based (survey) forecasts or model-based forecasts are what is often used for decision making for several reasons. Policy communication, anchoring of expectations, parameter calibration in structural or semi-structural models, and even policy rules, e.g., Taylor Rule, often depend on single estimates.

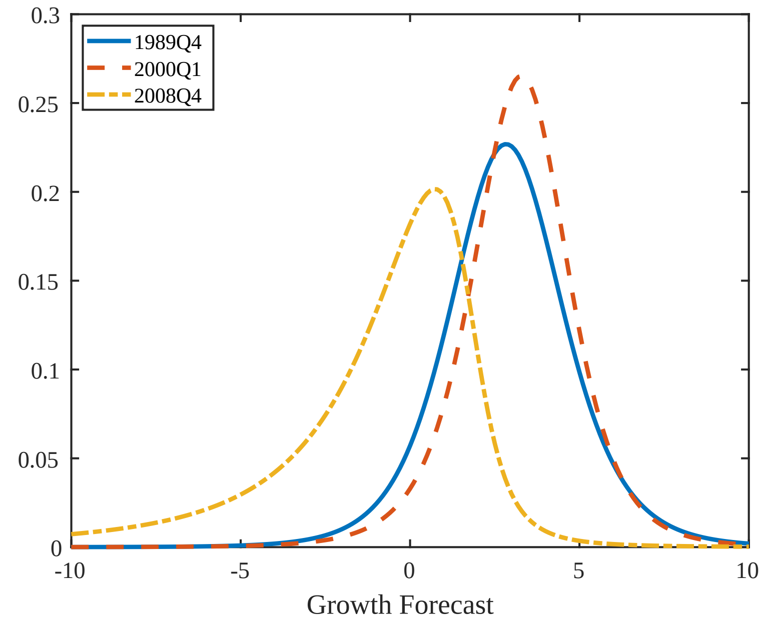
Uncertainty measurement around GDP growth conditional mean forecasts has become increasingly important too. From the fan charts introduced by the Bank of England in 1996, it is now a standard in various institutions including Bank of Canada, Norges Bank, the Federal Reserve Board of Governors, the NY FED, and the IMF. At first, this measurement was used to report confidence intervals or credible sets around the conditional mean (point) forecast providing a sense of precision in forecasting.

Nevertheless, the full density forecast has more information or predictive content than just the value of specific percentiles of the distribution. It includes information about its shape, dispersion, skewness, and other features from distributional outcomes. Economic and financial data often exhibit skewness, fat tails, or multimodal behavior. For example, when the economy faces extreme values, like booms or recessions, the growth rates should be represented in the tails, not close to the expected value.

Figure 1 shows selected dates of the one-year-ahead density forecasts for the U.S. GDP growth rate. They were computed using linear quantile regressions of the one-year-ahead real GDP growth on current real GDP growth and the National Financial Conditions Index (NFCI).[1] The resulting 5, 25, 75, and 95 percent quantiles were used to fit a skewed t-distribution as in Adrian et al. (2019).

[1]The target variable in the linear quantile regressions is the one-year-ahead year-over-year real GDP growth rate. However, the current real GDP growth rate used as covariate is an annualized quarter-over-quarter growth rate.

Figure 1: One-Year-Ahead U.S. GDP Growth Density Forecast



Note: The figure shows selected dates of the density forecasts computed using linear quantile regressions of the one-year-ahead real GDP growth on current real GDP growth and NFCI; then, the resulting quantiles are used to fit a skewed t-distribution as in Adrian et al. (2019).

It is evident that the shape of the density forecast changes through time. The forecast for 1989Q4 has a smaller average and larger variance relative to the forecast for 2000Q1. The explanation is that the 1989Q1 forecast comes after several years of tighten financial conditions and the 2000Q1 after years of loosen financial conditions. More striking is the forecast for 2008Q4 which accounts for the financial crisis. After near 15 years of loosen financial conditions, there was a sudden tightening starting at the end of 2007 which lead to the increased uncertainty for one-year-ahead forecast for 2008Q4. It is represented here with a large variance, smaller average, and large left tail. Overall, the time series of density forecasts could provide a way to improve conditional mean forecasts.

The jointly use of point and density forecasts is not straightforward, however. Point forecasts are scalar values whereas the density forecasts are infinite dimensional objects for continuous distributions, as it is the case for the real GDP growth.

In applications, density forecasts are made operationally feasible in several ways. Practitioners could use assumptions about the stochastic process, for example, a GARCH process with Gaussianity. They can compute density forecasts using Bayesian estimation or by fitting in some nonparametric procedure like kernel density estimation. Additionally, practitioners use a sufficiently large finite grid of values evaluated in a known parametric or nonparametric distribution. Still, how to use these densities in conjunction with the conditional mean forecast is not an easy task.

The literature has focused in forecasting combination for the joint usage of different predictive models, for example, Bayesian Model Averaging or Quantile Density Combination. These techniques do not exploit all the predictive content of the respective density forecast for obtaining conditional mean forecasts since the predictive content of the densities is summarized in single weights for each density. To the extend of our knowledge, whether the full predictive content of density forecasts are useful to improve conditional mean forecasts and, if applicable, how to implement this improvement are issues not been exploited yet in the literature.

In this paper, we propose a method to assess whether it is possible to improve a given conditional mean (point) forecast with a given density forecast by using the additional predictive content contained in the density forecast. It uses functional regression estimates of the h-period ahead target variable on the respective current conditional mean and density h-period ahead forecasts with current information for that target variable.

Since density forecasts for continuous variables are infinite dimensional objects which cannot be used directly in standard regression, we rely on functional principal components (FPCs) to effectively summarize and exploit the predictive content of density forecasts. We use the basis conformed by the leading FPCs of the sample variance operator of the density forecasts. It allows us to obtain an isometry between the Hilbert subspace of dimension K– of square integrable functions whose integral vanishes– and the real numbers of dimension K ($\mathbb{R}^K$) for the (weights) coordinates of the basis. Using these leading FPCs, we are able to effectively approximate our functional regression to a standard regression which can be estimated with OLS.

We perform an application to forecasting the U.S. quarterly GDP growth rate one-year-ahead with vintage data for real-time forecasting exercises and revised data for estimation and inference purposes.[2] We consider the Survey of Professional Forecasters (SPF)' consensus (average) forecast as a point forecast for our target variable.

Our baseline functional regression specification is what we call unrestricted. As a robustness check, we also use other functional regression specifications called by us as restricted, and centered with mean factor to study whether there is improvement and how the density forecast helps the point forecast. The unrestricted functional regression uses jointly the point and density forecast with agnostic weights. The restricted functional regression assigns full

[2]We also perform one-quarter-ahead forecasts as a robustness. Additionally, we use several parametric and nonparametric ways to compute density forecast with and without a data-rich environment. These results are included in the Supplementary Material as alternative results since they use only revised (last vintage) data due to lack of vintage data availability for several variables considered.

weight, i.e. equal to one, to the point forecast meaning the density forecast has a secondary role of explaining any systematic information left unexplained by the point forecast. We can also separate the density forecast in two components: the mean factor and the centered density. This allows us to know whether the mean factor or the rest of the density have a role in the improvement. We consider unrestricted and restricted variants with mean factor and centered density.

Our results show more than 30% improvement in the out-of-sample (OOS) root mean square (forecasting) error for one-year-ahead GDP growth by using our functional regression forecasts relative to the SPF. The main reason is the predictive content associated with recession and expansion periods which helps decreasing the OOS bias. Our findings suggest that the SPF's consensus (average) forecast tends to be conservative.[3]

Our paper is related with several growing branches of the literature starting with quantile regression and density forecasting in empirical macroeconomics (Carriero et al., 2024a,b; Corradi et al., 2023). Typically, conditional mean forecasts and density forecasts are prepared by distinctive working groups using different sets of information. Good predictors (covariates) and models for one are not necessarily good for the other. Model-based conditional mean forecasting could use regression methods like standard econometric time series models or machine learning algorithms. The conditional density of future real GDP growth can be obtained with Bayesian methods (Carriero et al., 2024a), or by fitting a density using quantile regression (Adrian et al., 2019; Carriero et al., 2024b; Lopez-Salido and Loria, 2024). The way of testing OOS forecasting performance also differs between density and point forecasts (Clark and McCracken, 2013; Corradi and Swanson, 2006; Corradi et al., 2023; West, 2006). Our paper offers a link between conditional mean forecasts and density forecasts, both for computation, combination, and performance testing purposes. It is possible since our proposed functional regressions, using density forecasts as covariates, bring point forecast estimates.

Our paper is also related to the growth-at-risk literature. It emphasizes the use of density forecast for studying its shape and defining the determinants of downside risk for economic growth. Deteriorating financial conditions are associated with an increase in downside risk for economic growth (Adrian et al., 2019, 2022). The main contributor of downside risk is a decline in investment (Amburgey and McCracken, 2023a). Our paper offers a way to align conditional mean forecasts with the associated downside or upside risks in the target variables.

Finally, the literature on functional regressions and functional time series and its use

[3]We do not find any improvement for one-quarter-ahead forecast since the information coming from the density forecast is concentrated in the mean factor, which is highly correlated with the point forecast.

in empirical macroeconomics and finance is also related to our work. (Bosq, 2000; Chang et al., 2016, 2021, 2024). This literature offers the theoretical backbone for our functional regressions. To the extend of our knowledge, our paper is the first to apply these type of regressions with FPCs basis to improve conditional mean forecasts with density forecasts using vintage and revised data for the U.S.

The outline of the paper is as follows. Section 2 details our econometric approach to measure the improvement of a given point forecast by using a given density forecast. It includes our baseline functional regression specification used and the estimation methodology. Section 3 describes the data we use for our application to forecasting the U.S. quarterly GDP growth. Section 4 explains the forecasting methodology, how we evaluate the out-of-sample forecasting performance, and how we perform inference to explain how the density forecast helps in the forecasting improvement. Section 5 presents our estimation and inference results with our baseline functional regression specification using revised data. Section 6 shows results using vintage data for a real-time forecasting environment. Section 7 presents several robustness checks and alternative results. Finally, in Section 8 we provide some concluding remarks.

2 Econometric Methodology

2.1 The Model

Our goal is to assess whether it is possible to improve the conditional mean (point) forecast of a target variable by using additional predictive content contained in a respective density forecast. Let $f_{Y_{t+h}|X_t} \equiv f_{t+h|t}$ represent the density forecast computed at time t for the target variable, y, to be observed at time $t+h$ with forecasting horizon h, and information of covariates (predictors), x, up to time t.[4] Let $\left(f_{t+h|t}\right)_{t=1}^{T-h}$ be the complete time series of density forecasts obtained in some way. It is a proper functional time series. Its sample variance operator, called $\bar{\Gamma}$, is formally defined in a Hilbert space H of square integrable functions whose integral vanishes.

In this Hilbert space, the inner product is given by

$$\langle u, v \rangle = \int u(r)v(r)dr,$$

[4]We are using the notation $f_{t+h|t}$ to represent $f_{Y_{t+h}|X_t}$ for simplicity in this section. It means that $f_{t+h|t}$ is formally a function representing the conditional density $f_{Y_{t+h}|X_t}\left(y_{t+h}|x_t\right)$ with X_t the available information set potentially including the current realization at t of the target variable, Y_{t+h}.

for $u, v \in H$, and the norm given by

$$\|v\|^2 = \langle v, v \rangle = \int v^2(r)dr,$$

for $v \in H$.

The improvement from the density forecast could be measured with the following functional regression

$$(y_{t+h} - \bar{y}) = \alpha \left(m_{t+h|t} - \bar{m}\right) + \left\langle \beta, \left(f_{t+h|t} - \bar{f}\right) \right\rangle + \varepsilon_{t+h} \tag{1}$$

where (y_{t+h}) is the realization of the time series for the target variable indexed at time $t+h$, $\left(m_{t+h|t}\right)$ is the point forecast for (y_{t+h}) conditional of information up to t, and $\left(f_{t+h|t}\right)$ is the density forecast for (y_{t+h}) conditional of information up to t. $\bar{y}$, $\bar{m}$, and $\bar{f}$ are sample averages of $(y_{t+h})_{t=1}^{T-h}$, $\left(m_{t+h|t}\right)_{t=1}^{T-h}$, and $\left(f_{t+h|t}\right)_{t=1}^{T-h}$, respectively.[5] h is the forecasting horizon meaning that the one-quarter and one-year ahead forecasts with quarterly data are represented by $h = 1$ and $h = 4$, respectively. α is a scalar parameter and β is a functional parameter.

Equivalently, we may define the functional regression in equation (1) as

$$y_{t+h} = \tau + \alpha\, m_{t+h|t} + \left\langle \beta, f_{t+h|t} \right\rangle + \varepsilon_t, \tag{2}$$

with an additional constant term $\tau = \bar{y} - \alpha\bar{m} - \left\langle \beta, \bar{f} \right\rangle$. Equation (1) is used for estimation purposes. The details for estimation are explained in Subsection 2.2. However, forecasting and performance evaluations use equation (2), since its left-hand-side includes only (y_{t+h}). Its fitted values, $(\hat{y}_{t+h})$, are point forecasts themselves, with an appropriately recovered τ coefficient from the estimations. We consider several variations of equation (2) as a robustness check which will be explained in Section 7.

By the Riesz representation theorem, any linear functional $\ell : H \to \mathbb{R}$ of $\left(f_{t+h|t}\right)$ can be represented as $\ell\left(f_{t+h|t}\right) = \left\langle \beta, f_{t+h|t} \right\rangle$ for some $\beta \in H$. Therefore, our functional regression model captures any linear effect of $\left(f_{t+h|t}\right)$ on (y_{t+h}) with the functional coefficient β.

Note that $\bar{f}$ and each $\left(f_{t+h|t}\right)_{t=1}^{T-h}$, for each t and h, are proper densities. Hence, we have

$$\int \left(f_{t+h|t} - \bar{f}\right)(r)dr = 0$$

[5]We consider the sample mean of $\left(f_{t+h|t}\right)_{t=1}^{T-h}$, computed as $\bar{f} = \frac{1}{T-h}\sum_{t=1}^{T-h} f_{t+h|t}$, to estimate the population mean, $\mathbb{E}f_{t+h|t}$, assumed to be common for all $t = 1, \ldots, T-h$.

for all t. This implies that $f_{t+h|t} - \bar{f} \in H$, from which it follows that

$$\int \beta(r)dr = 0$$

since $\beta \in H$. The functional regression coefficient β is thus given as a function whose integral vanishes.

2.1.1 Specifications to Estimate

We investigate whether using density forecast improves out-of-sample (OOS) forecasting performance of the point forecast, $(m_{t+h|t})$, by estimating equation (1). We consider the following specification:

$$y_{t+h} = \tau + \alpha m_{t+h|t} + \langle \beta, f_{t+h|t} \rangle + \varepsilon_{t+h} \tag{ur}$$

which we call an *unrestricted functional regression* or *ur* for short. It measures the jointly use of $(m_{t+h|t})$ and $(f_{t+h|t})$ to forecast (y_{t+h}) by giving agnostic weights. The resulting fitted values, $(\hat{y}_{t+h|t})$, from the functional regression estimation in equation (*ur*) can be used as an alternative point forecast.

An estimate of the functional coefficient $\hat{\beta}$ indicates the effect of the density forecasts on the target variable. $\hat{\beta}$ is a function with domain in $\mathbb{R}$ representing possible point forecast values in the support of $(f_{t+h|t})$, and with codomain in $\mathbb{R}$ giving us suggested adjustments to those values. When we discuss $\hat{\beta}$ values, we are referring to $\hat{\beta}$ codomain values. For example, statistically significant positives values (in the codomain) tell us that the point forecast (in the domain) should be adjusted upwards to improve its forecasting performance given the information from $(f_{t+h|t})$.

If the estimated functional coefficient $\hat{\beta}$ is statistically different from zero, then, the density forecasts, $(f_{t+h|t})$, predict the target variable beyond the predictive content of the point forecast. It is a direct implication of the Frisch–Waugh–Lovell theorem since our functional regressions are linear regression specifications. Therefore, $\hat{\beta}$ tells us a way of aligning the point forecasts with the underlying distributional risks or uncertainties measured with the density forecasts.

The specification in equation (*ur*) does not assume the point forecast is correct on average. Considering the Frisch-Waugh-Lovell Theorem, it allows the density forecast to additionally suggest adjustments related to anything unexplained by the point forecast in approximating the expected value or any other region of the (unknown) density of the target variable. Furthermore, the specification in equation (*ur*) is flexible enough to assign any weight in the real numbers to the point forecast.

As a robustness check, we also use other functional regression specifications called by us as restricted, and centered with mean factor to study how the density forecast provides forecasting improvement. These will be detailed in Section 7.

2.2 Estimation Methodology

In equation (1), the component $\langle \beta, (f_{t+h|t} - \bar{f}) \rangle$ includes the coefficient β and the covariate $(f_{t+h|t} - \bar{f})$ which are infinite-dimensional functionals. They cannot be used directly in the estimation. Nevertheless, it is feasible to estimate equation (1) by following the estimation procedure proposed here.

Let the sample variance operator of $(f_{t+h|t})_{t=1}^{T-h}$ be

$$\bar{\Gamma} = \frac{1}{T-h} \sum_{t=1}^{T-h} \left((f_{t+h|t} - \bar{f}) \otimes (f_{t+h|t} - \bar{f}) \right) \tag{3}$$

It is defined in the Hilbert space of square integrable functions whose integral vanishes, H, with $\otimes$ the tensor product.[6]

By the Parseval's identity, any $v \in H$ can be written as

$$v = \sum_{k=1}^{\infty} \langle u_k, v \rangle u_k$$

using any countably infinite orthonormal basis (u_k). Each u_k is also a function and $u_k \in H$. We can select an appropriate orthonormal basis for estimation purposes.

A finite-dimensional representation of $\bar{\Gamma}$ is required to estimate equation (1). We use the (normalized) FPCs of $(f_{t+h|t})_{t=1}^{T-h}$, which are the eigen-functions of $\bar{\Gamma}$ in equation (3), denoted by $(v_k)_{k=1}^{T-h}$, and associated with eigenvalues $(\lambda)_{k=1}^{T-h}$, $\lambda_1 > \cdots > \lambda_{T-h}$. (v_k) is our chosen countably infinite orthonormal basis.

Therefore, we can approximate $f_{t+h|t} - \bar{f}$ as

$$f_{t+h|t} - \bar{f} \approx \sum_{k=1}^{K} \langle v_k, f_{t+h|t} - \bar{f} \rangle v_k \tag{4}$$

for $t = 1, \ldots, T-h$ with an appropriately chosen value of K, the finite number of FPCs,

[6]The tensor product, $\otimes$, in the Hilbert space, H, corresponds to the outer product in $\mathbb{R}^n$. In fact, if $u, v \in \mathbb{R}^n, u \otimes v = uv'$, which contrasts with the inner product $\langle u, v \rangle = u'v$. In a general Hilbert space H with inner product $\langle \cdot \cdot \rangle$, the tensor product $u \otimes v$ for any $u, v \in H$ yields an operator of rank one, which is formally defined as $(u \otimes v)w = \langle v, w \rangle u$ for all $w \in H$.

such that $1 \leqslant K \leqslant T - h$.

Our approximation in equation (4) is based on two important choices: the choice of the basis (v_k) and the choice of the truncation number K. We follow the standard practice of the functional data analysis (FDA) and use the leading eigen-functions (v_k) of the sample variance operator $\bar{\Gamma}$ in equation (3). We will refer to them simply as the FPC basis. The reader is referred to Chang et al. (2023) for some optimal properties of using the FPCs as a basis in this context.

The most effective way to represent the temporal variation of a functional covariate in functional regressions is by using the FPC basis (Chang et al., 2023). We choose the truncation number K with Bayesian Information Criteria (BIC). We obtain a K value for our functional regression equation (*ur*). Subsection 4 provides more details about the empirical computation of the K value. Given a truncation number K, the identification of the functional regressions works as follows.

Let V be the K-dimensional subspace of H spanned by $(v_k)_{k=1}^K$, and define Π to be the (orthogonal) projection on V in H. Then it follows that

$$\Pi\left(f_{t+h|t} - \bar{f}\right) = \sum_{k=1}^{K} \left\langle v_k, f_{t+h|t} - \bar{f}\right\rangle v_k,$$

which we may represent as K-dimensional vectors in the coordinate system given by $(v_k)_{k=1}^K$.

For precision, we introduce a mapping π on H defined as $\pi(v) = ⦇v⦈$, where

$$⦇v⦈ = \begin{pmatrix} \langle v_1, v\rangle \\ \vdots \\ \langle v_K, v\rangle \end{pmatrix}, \tag{5}$$

for any $v \in H$. For any $v \in V$, there exists a unique K-dimensional vector $⦇v⦈$, and we have that

$$\|v\|^2 = \sum_{k=1}^{K} \langle v_k, v\rangle^2 = \|⦇v⦈\|^2,$$

where we use the notation $\|\cdot\|$ to denote the Hilbert space norm in H for $v \in V$ on the left-hand side and the usual Euclidean norm for $⦇v⦈$ in $\mathbb{R}^K$ on the right-hand side. Therefore, the mapping π defines an isometry between V and $\mathbb{R}^K$, which are identical, for the coordinates $\langle v_k, v\rangle$. This allows us to represent the functional time series $\left(f_{t+h|t} - \bar{f}\right)$ effectively as the K-dimensional vector time series $\left(⦇f_{t+h|t} - \bar{f}⦈\right)$, i.e. as K vectors each one of size $T - h$ representing the loadings of the FPCs.

Now, we consider the functional term in our functional regression in equation (1), which we may write as

$$\left\langle \beta, \left(f_{t+h|t} - \bar{f}\right)\right\rangle = \left\langle \beta, \Pi\left(f_{t+h|t} - \bar{f}\right)\right\rangle + \left\langle \beta, (1-\Pi)\left(f_{t+h|t} - \bar{f}\right)\right\rangle \approx \left\langle \beta, \Pi\left(f_{t+h|t} - \bar{f}\right)\right\rangle,$$

where $(1-\Pi)(\cdot)$ represents the orthogonal projection on $V^{\complement}$ in H. Since $f_{t+h|t} - \bar{f} \notin V^{\complement}$, the component $\left\langle \beta, (1-\Pi)\left(f_{t+h|t} - \bar{f}\right)\right\rangle$ could be ignored.

Moreover, we have

$$\begin{aligned}
\left\langle \beta, \Pi\left(f_{t+h|t} - \bar{f}\right)\right\rangle &= \left\langle \beta, \sum_{k=1}^{K} \left\langle v_k, \left(f_{t+h|t} - \bar{f}\right)\right\rangle v_k \right\rangle \\
&= \sum_{k=1}^{K} \langle v_k, \beta\rangle \left\langle v_k, \left(f_{t+h|t} - \bar{f}\right)\right\rangle \\
&= (\beta)'(f_{t+h|t} - \bar{f}),
\end{aligned}$$

and therefore, our functional regression in equation (1) reduces to

$$(y_{t+h} - \bar{y}) \approx \alpha\left(m_{t+h|t} - \bar{m}\right) + (\beta)'(f_{t+h|t} - \bar{f}) + \varepsilon_{t+h}. \tag{6}$$

If we use the approximation for $\left(f_{t+h|t} - \bar{f}\right)$ introduced in equation (4), then, the regression in equation (6) is entirely standard having no functional arguments, and the parameters $\alpha \in \mathbb{R}$ and $(\beta) \in \mathbb{R}^K$ can be estimated by the usual OLS method. Furthermore, we can estimate the functional regression coefficient β by

$$\hat{\beta} = \pi^{-1}(\widehat{(\beta)}), \tag{7}$$

where π^{-1} is the inverse of the isometry π introduced in equation (5), and $\widehat{(\beta)}$ is the OLS estimate obtained from the regression in equation (6).

Since the estimating regression in equation (6) is a standard regression, we can use the standard wild bootstrap procedure for regressions to obtain uncertainty measurements and for inference purposes. Specifically, we construct the artificial samples and bootstrap critical values using the procedure outlined in Appendix A.

2.3 Functional Regression Intuition

The sample variance operator, $\bar{\Gamma}$, in equation 3 should be regarded not as a matrix, but as a hypercube. Like a cube, some dimensions are related to the length, others to the width, and the height. For $\bar{\Gamma}$, these are the time dimension, the domain, and the codomain of

$(f_{t+h|t})_{t=1}^{T-h}$.

In summary, we can approximate $f_{t+h|t} - \bar{f}$ as a sum of coordinates (weights), $\langle v_k, f_{t+h|t} - \bar{f} \rangle$, times the respective FPCs or (eigen)functions v_k. However, we only pick K FPCs, with $K < T-h$, and we call the resulting subspace V. Everything in V is a function, including its basis, the selected FPCs. The reward we get from this is that these functions are orthogonal and countable, no matter how complex the original function was.

We map (through an isometry) V to $\mathbb{R}^K$, but this mapping refers only to the coordinates needed for the basis to reconstruct the sample variance operator. In other words, what we obtain in $\mathbb{R}^K$ is a vector of coordinates for each FPC to span the sample variance operator. For example, if we pick $K = 1$, we are using only the leading FPC (which is still a function), and we need only one coordinate. This coordinate is not the factor loading but the coefficient associated with the basis function that spans the sample variance operator.

At the end, we have a real vector of coordinates (of dimension K) and discretized functions (each FPC). Together, they represent (summarize) the information of the original sample variance operator and, hence, of the (demeaned) density forecast time series. When we say that the functional time series $(f_{t+h|t} - \bar{f})$ is effectively represented as a K-dimensional vector time series $((f_{t+h|t} - \bar{f}))$, we are referring to $((f_{t+h|t} - \bar{f}))$ being a $(T-h) \times K$ matrix. If we have $K = 1$, then $((f_{t+h|t} - \bar{f}))$ is a vector of dimension $T - h$, and we can use it for estimation with OLS.

Of course, both the density forecasts and the FPCs are functions. We cannot use an infinite set of numbers as support for the density forecast or as domain for the FPCs. We discretize the densities, and by construction, the sample variance operator and the respective FPCs follow the same discretization. Discretization is not critical for the rest of the results as long as it meaningfully represents the support of the density forecast. The only difference lies in how much granularity we wish to allow for the density forecasts and the FPCs. This discretization is valid since the mapping π has allowed us to work with functions in the real numbers, $\mathbb{R}$, instead of objects in the Hilbert space.

Let us denote the number of discrete elements by W. Then, each density is represented by an $W \times 1$ vector of probabilities in its codomain associated with support values in its domain which could be regarded as labels. Since we have $T-h$ densities, the complete set of demeaned densities is stored in a $(T-h) \times W$ matrix. The resulting sample variance operator has dimension $W \times W$. Each FPC is a $W \times 1$ vector –the discretized eigen-function (similar to an eigenvector)– with labels the chosen W support values of the density. Multiplying the $(T-h) \times W$ matrix of demeaned densities by a given FPC yields the corresponding factor loading vector, which has dimension $(T-h) \times 1$.

The respective factor loadings for the selected FPCs are those we include in the OLS estimation. Therefore, the matrix of covariates in the standard (vector-based) regression—equivalent to the functional regression—has dimension $(T-h) \times (K+L)$, where K is the number of selected FPCs and L represents the number of additional covariates. In our case, $L=2$ if the regression includes an intercept and the point forecast, and $L=3$ if we also include the mean factor, which is already a vector by construction. In our applications, we take a sequence beginning in -20 and finishing in 20 with jumps of size 0.1 for a total of 401 support values for the density forecast, i.e. we use $W=401$. Here 20 means a growth of 20%.[7]

3 Data

Our target variable, denoted by y_{t+h} with h the forecasting horizon, is the year-over-year quarterly real GDP growth rate for one-year-ahead forecasting ($h=4$). We choose the year-over-year rates for the one-year-ahead forecasting exercise instead of the commonly used annualized quarter-over-quarter rates because year-over-year growth rates are smoother. We also use annualized quarter-over-quarter real GDP growth rate for one-quarter-ahead forecasting ($h=1$) as a robustness check.

Our sample period is 1973Q1-2024Q3. We focus on data for the United States to obtain a large time series sample for the point forecast and several covariates.[8] We use the average (consensus) forecast from the Survey of Professional Forecasters as a benchmark point forecast. More details about the SPF are provided in Appendix B.

To compute conditional quantiles and the density forecasts, we consider several covariates suggested from the literature. Financial conditions are a determinant of downside risk for real GDP growth (Adrian et al., 2019, 2022). Thus, our baseline model covariates (x_t) for real GDP growth include the current values of the National Financial Conditions Index (NFCI) and the current values of the annualized quarter-over-quarter real GDP growth rate for both $h=1$ and $h=4$.[9]

[7]It is also possible to do a selective discretization of the density forecasts and hence of the FPCs. In computational economics, it is common to give higher weights to certain sections of the domain. We assigned the same weights to all sections of the density forecast and let the estimated probabilities determine the likelihood of observing the respective values of growth rates.

[8]We will use the terminology predictors or covariates interchangeably in the paper.

[9]Additional specifications using 271 macroeconomic variables are also used, and we refer to those as data-rich environment. Nonetheless, it was not possible to obtain a real-time dataset for these data-rich environment specifications because several of the variables do not have enough vintages available. We still consider these specifications for conditional quantiles to study which variables have predictive content in the different regions of the density forecasts. Baseline forecasts using revised data will be comparable to these other specifications. These specifications suffer from a forward-looking bias since their estimation uses revised data. More details about these specifications are showed in the Supplementary Material.

We use both vintage and revised (last vintage available) data.[10] We use revised data for estimation and inference purposes whereas the vintage data is used for a real-time forecasting exercise.[11]

All data is obtained from the FRED or ALFRED Datasets except the point forecasts obtained from the SPF.[12] Oldest historical vintages for the NFCI are taken from Amburgey and McCracken (2023b) with first observation available for 1973Q1.

4 Estimation, Inference, and Forecasting Methodology

We perform two main exercises to assess whether the jointly use of density and point forecast provides improvement for forecasting. First, we use revised data for estimation and inference purposes to explain how the density forecast provides forecasting improvement, if there is any. Those results are shown in Section 5. Second, we use vintage data in a real-time forecasting environment to measure whether there is actual forecasting improvement. Those results are shown in Section 6.

This section outlines the procedures used to generate out-of-sample (OOS) quantile forecasts and density forecasts, estimate functional regressions, evaluate forecasting performance, and perform inference. The overall steps for forecasting are: i) compute the quantile forecasts; ii) fit or estimate density forecasts; iii) obtain the functional principal components (FPCs); iv) choose the number of FPCs; v) estimate the functional regressions and obtain their forecasts (fitted values); and vi) perform OOS forecast evaluation for the point forecast and the proposed functional regressions. The steps for inference involve i) to v) using revised instead of vintage data, and using the full sample for estimation instead of a rolling window.

Sections 4.1 and 4.2 explains the quantile and density forecasting procedures, respectively. Section 4.3 details the empirical estimation of our functional regression and the computation

[10]Vintage data refers to the first release of the previous quarter data. For example, the level of real GDP for 1965Q3 is reported by the Bureau of Economic Analysis (BEA) at the end of October 1965, i.e., 1965Q4 and hence the vintage data references 1965Q4 with last observed value for 1965Q3. The BEA revises that number for the next two months (end of November and December) before giving a final revised value. This means there exist other 1965Q4 vintages of data – which we are not using. We follow the standard practice in the forecasting literature of using the first vintage of the quarter to represent the quarter. For simplicity, we will refer to the date of the observed values and not the date of the vintage when discussing sample dates, estimation, and out-of-sample performance evaluation.

[11]The forecasting exercise with vintage data is usually called "real-time" forecasting since the estimations are performed as if a forecaster is computing them every time when new observations are available. Using the last vintage of the data, i.e., revised data, is often called "pseudo-real-time" forecasting as it does include the information from data revisions. A fair comparison between the point forecasting and our proposed method should use the information set available to survey respondents which coincides with a real-time forecasting exercise as explained in Appendix B for the SPF.

[12]When using a data-rich environment exercise, the S&P, NYSE, and NASDAQ stock price indexes are gathered from Yahoo Finance.

of their forecasts. Section 4.4 explains how we evaluate the forecasting performance of our models. Section 4.5 provides more details about our functional regression inference exercise.

4.1 Quantile Forecasting

We compute one-quarter-ahead, $h = 1$, and one-year-ahead, $h = 4$, OOS quantile forecasts using a rolling window approach. Specifically, we estimate the conditional linear quantile functions of future real GDP growth, y_{t+h}, given covariates x_t, for various quantile levels $\varrho \in (0, 1)$.

The coefficients are estimated using the quantile loss function. Namely,

$$\hat{Q}_{\varrho}\left(y_{t+h}|x_t\right) = x_t'\hat{\gamma}_{\varrho} \tag{8}$$

with

$$\hat{\gamma}_{\varrho} = \underset{\gamma_{\varrho}\in\mathbb{R}^M}{\arg\min} \sum_{t=1}^{T-h} \left(\varrho \cdot \mathbf{1}_{(y_{t+h}\geqslant x_t'\gamma_{\varrho})} \left|y_{t+h} - x_t'\gamma_{\varrho}\right| + (1-\varrho)\cdot \mathbf{1}_{(y_{t+h}<x_t'\gamma_{\varrho})} \left|y_{t+h} - x_t'\gamma_{\varrho}\right|\right) \tag{9}$$

where $\mathbf{1}_{(\cdot)}$ denotes the indicator function, y_{t+h} is the target variable for period $t+h$, x_t is a vector of M covariates up to time t, and $\hat{\gamma}_{\varrho}$ is the vector of estimated coefficients for quantile ϱ. Koenker and Bassett (1978) show consistency of $\hat{Q}_{\varrho}\left(y_{t+h}|x_t\right)$ for the linear estimator in equation (8) with the coefficient estimate obtained from equation (9).

For the real GDP growth as a target variable in period $t + h$, we estimate the linear quantile specifications using as covariates the current values of financial conditions measured with the NFCI, as in Adrian et al. (2019), together with the current values of real GDP growth.

Our baseline rolling window begins in 2004Q1. We use data from 1973Q1 to 2003Q4 and from 1973Q1 to 2003Q1 to fit the models for $h = 1$ and $h = 4$, respectively. Then, they are used to forecast the quantiles of the target variable for 2004Q1. Afterwards, the window advances by one quarter. This procedure is repeated sequentially and produces a full set of rolling quantile forecasts until the last forecast is computed for 2024Q3 with a total of 83 OOS points.

4.2 Density Forecasting

We estimate few quantiles and use the resulting quantile function, $\hat{Q}_{\varrho}\left(y_{t+h}|x_t\right)$, giving us quantile forecasts to compute density forecasts by fitting a skewed t-distribution developed by Azzalini and Capitanio (2003).[13] The skewed t-distribution is the following:

$$f(y;\mu,\sigma,\alpha,\nu)=\frac{2}{\sigma}t\left(\frac{y-\mu}{\sigma};\nu\right)T\left(\alpha\frac{y-\mu}{\sigma}\sqrt{\frac{\nu+1}{\nu+\left(\frac{y-\mu}{\sigma}\right)^2}};\nu+1\right)$$

with $t(\cdot)$ and $T(\cdot)$ denoting the probability density function and the cumulative distribution function of the Student t-distribution, respectively. μ is the location, σ is the scale, ν is the fatness, and α is the shape; all parameters of the skewed t-distribution.

For each quarter, the density forecast fitting process requires finding $\{\hat{\mu}_{t+h},\hat{\sigma}_{t+h},\hat{\alpha}_{t+h},\hat{\nu}_{t+h}\}$ as follows:

$$\{\hat{\mu}_{t+h},\hat{\sigma}_{t+h},\hat{\alpha}_{t+h},\hat{\nu}_{t+h}\}=\underset{\mu,\sigma,\alpha,\nu}{\arg\min}\sum_{u}\left(\hat{Q}_{\varrho}\left(y_{t+h}\mid x_t\right)-F^{-1}(u;\mu,\sigma,\alpha,\nu)\right)^2$$

where $F^{-1}(u;\mu,\sigma,\alpha,\nu)$ is the quantile function of the skewed t-distribution for given values of the μ_{t+h}, σ_{t+h}, α_{t+h}, and ν_{t+h} parameters. The minimization matches the 5%, 25%, 75%, and 95% percent quantiles of the estimated quantile forecasts with those of the skewed t-distribution by doing a numerical search for the parameter values that minimize the squared difference. As a result, we obtain a density forecast for each quarter and the corresponding time series is a proper functional time series.[14]

For in-sample periods in each rolling window, we fit the densities using the corresponding in-sample quantile estimates one period at the time. For the out-of-sample period, we estimate the predictive densities based on the rolling OOS quantile forecasts, similarly, one period at the time. The estimated density functions in-sample are used as inputs for the functional principal component analysis and the functional regression estimation. The OOS estimated density functions are used as inputs for the functional regression forecasts which can be considered as rolling window forecasts.

We evaluate the OOS accuracy and calibration of the density forecasts to assess their reliability. The predictive score is computed as the predictive distribution generated by a model and evaluated at the realized value of the time series. Higher predictive scores indicate more accurate density forecasts because the ex post realization is considered more likely. It is a relative performance measure between competing models.

[13]Estimates of our linear quantile specifications suffer from quantile crossing. By fitting a parametric distribution using few well-separated quantiles, we avoid problems from the crossing quantiles and we are able to approximate a density even when few quantiles are used. We select the skewed t-distribution since it has been useful for measuring tail risks (Adrian et al., 2019).

[14]The fitting density procedure outlined here could be regarded as a parametric approach. Some nonparametric approaches could also be used. The Supplementary Material details the parametric and the nonparametric approaches considered for density estimation in alternative exercises and robustness checks.

We also compute the empirical cumulative distribution of the probability integral transforms (PITs) of each density forecast. It measures the percentage of observations that are below any given quantile. The model is better calibrated the closer the empirical cumulative distribution of the PITs is to the 45-degree line. In fact, in a perfectly calibrated model, the cumulative distribution of the PITs is a 45-degree line. It means that the fraction of realizations below any given quantile $Q_\varrho(y_{t+h}|x_t)$ of the predictive distribution is exactly equal to ϱ. We report confidence bands around the 45-degree line to account for sample uncertainty following Rossi and Sekhposyan (2019).

4.3 Functional Regression Estimation and Forecasting

We estimate functional regressions after obtaining the sequence of density forecasts functions. A key step in this procedure is selecting the number K of FPCs to retain. We report scree plots to know how much information each FPC contains. Scree plots report the proportions of the cumulative temporal variations of the density forecast time series, $\left(f_{Y_{t+h}|X_t}\right)$, explained by leading FPCs.

We estimate the functional regressions for values of $K \in \{1, 2, \ldots, 5\}$.[15] For each value of K, we estimate the functional regression coefficients—intercept τ, slope α of the point forecast $\left(m_{t+h|t}\right)$, and the functional coefficient β of the density forecast—based on the full in-sample data in the first rolling window. It means we have several competing models. Then, we choose the truncation number K with Bayesian Information Criteria (BIC). This procedure ensures we fulfill the assumptions of Hansen et al. (2011) and we can obtain a K value for our functional regression equation (ur). Once K is fixed, we produce the OOS forecasts for the period 2004Q1-2024Q3 by using the OOS point forecast values, $\left(m_{t+h|t}\right)$, and OOS density forecasts, $\left(f_{Y_{t+h}|X_t}\right)$, and estimating the functional regression coefficients in each rolling window.

4.4 Forecast Evaluation

We evaluate the OOS forecast performance over the period 2004Q1–2024Q3, a total of 83 quarterly observations. The benchmark is the SPF mean forecast $m_{t+h|t}$. We compare its performance against forecasts generated by our functional regression specification ur.

We compute the bias, variance, and root mean squared (forecasting) error (RMSE) relative to those of the SPF. Additionally, we conduct Diebold-Mariano tests to assess whether the forecast improvements relative to the benchmark $m_{t+h|t}$ are statistically significant. The null hypothesis in each test is that the benchmark (SPF) and the competing model have

[15] Results from the scree plots tell us that 5 FPCs explain more than 90% of the temporal variation.

equal predictive accuracy. Negative numbers in the statistic tell us whether there has been an improvement.

The standard Diebold-Mariano (DM) statistic is the following (Diebold and Mariano, 1995):

$$DM_{12} = \frac{\bar{d}_{12}}{\widehat{\sigma}_{\bar{d}_{12}}}$$

where $DM \xrightarrow{d} N(0,1)$ under the null hypothesis $DM_{12} = 0$. $d_{12,t+h} = \varepsilon^2_{1,t+h} - \varepsilon^2_{2,t+h}$ is the difference in the squared forecasting errors of our proposed functional regression, $\varepsilon^2_{1,t+h}$, with respect to those of $m_{t+h|t}$, $\varepsilon^2_{2,t+h}$. $\bar{d}_{12}$ and $\widehat{\sigma}_{\bar{d}_{12}}$ are the sample average and standard deviation, respectively, of the difference in these forecasting errors.

4.5 Functional Regression Inference

For inference purposes only, not for forecasting, we use our full sample, 1973Q1-2024Q3, with revised data. The goal is to summarize the scree plots, density scores, and PITs. We also report the shape of the leading FPCs, the magnitude, and statistical significance of the functional regression estimated coefficients, and discuss their interpretation. We also estimate here the functional regressions for values of $K \in \{1, 2, \ldots, 5\}$ and choose the truncation number K with BIC using the full sample. Inference is performed using the bootstrapped confidence intervals resulting from the procedure outlined in Appendix A.

5 Estimation and Inference Results

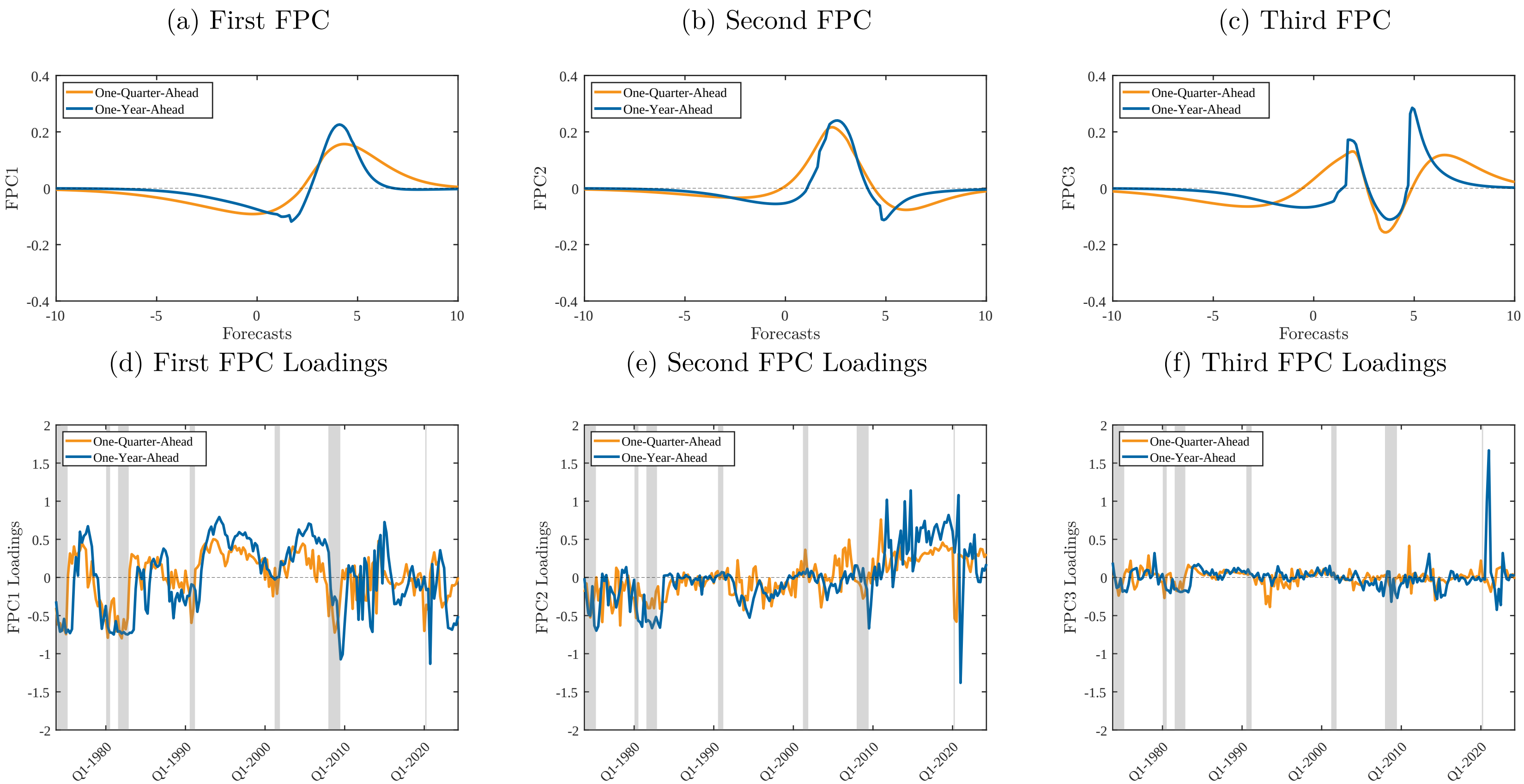
This Section shows our estimation and inference results using our full sample, 1973Q1-2024Q3, with revised data. We summarize our findings related to the density forecast and Functional Principal Component Analysis, as well as our functional regression estimates and their estimated coefficients interpretation.

5.1 Density Forecasts and FPCA

Figure 2 shows the first three leading FPCs and their respective loadings for the GDP growth density forecast time series. We call interchangeably FPCs as factors for simplicity.

The loadings associated with a factor are a time series. We may interpret them as the time series of coefficients for the factor measuring its effect at each time. Factors are identified up to a sign meaning that a negative loading is equivalent to transform it to a positive loading if we flip the respective factor. Figure 2 shows that factors and theirs loadings have a similar interpretation between one-quarter-ahead and one-year-ahead forecasts for GDP growth.

Figure 2: Functional Principal Components and Their Loadings



Note: The figure shows the first three leading functional principal components and their loadings for the respective density forecasts with sample 1973Q1-2024Q3. The density forecasts are computed using linear quantile regressions of the one-year-ahead real GDP growth on current real GDP growth and NFCI; then, the resulting quantiles are used to fit a skewed t-distribution. The gray rectangles represent U.S. recession dates as classified by the NBER.

Take as a reference the first factor and its loadings for GDP growth one-year-ahead forecasts shown in Figure 2, Panels (a) and (d), respectively. When the leading factor function intercepts zero in the y-axis (see its shape in Panel (a)), the x-axis has a value of around 2.5%. It means that when the leading factor loading is positive (see the loadings in Panel (d)), the distribution of forecasts is being moved to the right by suppressing smaller forecasts less than about 2.5% while increasing the occurrence of larger forecasts above 2.5%. In Panel (d), the gray rectangles represent recession dates as classified by the National Bureau of Economic Research (NBER). It shows that the first factor loadings are related with recession periods when the loading goes down, and expansion periods, when the loading goes up.

The loadings of the first factor and the point forecasts, $(m_{t+h|t})$, are positively correlated with the sample correlation coefficients 0.4714 and 0.4721 for GDP growth one-quarter-ahead and one-year-ahead forecasts, respectively. Therefore, the point forecast tends to increase when the loading of the first factor gets larger, as is well expected given the larger weight for large forecasts.

We report up to three leading factors and their loadings in Figure 2. We do not report other factors and their loadings, since they seem to be largely uninterpretable and unrelated to the point forecasts. Overall, our results suggest that the first factor has the most predictive content and interpretability. Functional regression specifications with only one factor were preferred due to its good forecasting performance as well as their interpretability.

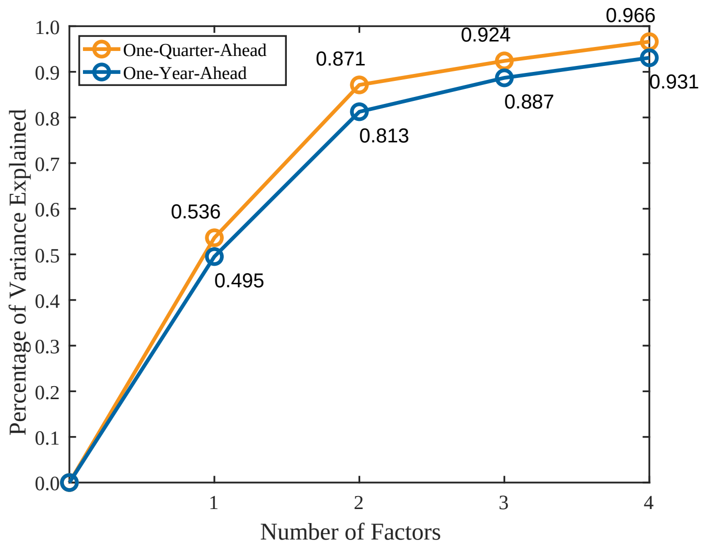
Figure 3 shows the scree plots. They report the cumulative percentage of the variance of the respective density forecast time series explained by the leading factors. The first leading FPC explains about 53.6% for one-quarter-ahead forecasts ($h = 1$) and 49.5% for one-year-ahead forecasts ($h = 4$). Similarly, the first and second FPCs explain jointly about 87.1% and 81.3%, for $h = 1$ and $h = 4$, respectively. The four leading FPCs explain more than 90% of the variation.

5.1.1 Out-of-Sample Accuracy: Scores and PITs

We assess the reliability of the density forecasts with measures of their accuracy and their calibration. Figure 4 shows the predictive scores, computed as the predictive distribution generated by a model and evaluated at the realized value of the time series in Panel (a).

Predictive scores are larger for one-year-ahead forecast relative to one-quarter-ahead forecast. Recession times are reported with gray rectangles and it is evident that the scores fall down during recession times. Even though, forecasting recession dates is of interest, it is a difficult task since there is few information or data available for these rare events from the perspective of the density forecast. It makes sense that the scores are low there since they

Figure 3: Scree Plots



Note: Figure shows the cumulative percentage of the variance of the respective density forecasts explained by the leading functional principal components (factors).

measure the likelihood of an event; being recession and expansion periods rare events, they should be located in the tails of the distribution, i.e. with low likelihood by construction. In general, the predictive scores shown here are similar to others presented in the literature, like those of Adrian et al. (2019) for GDP growth, meaning that our density forecasts are reliable from an accuracy perspective.

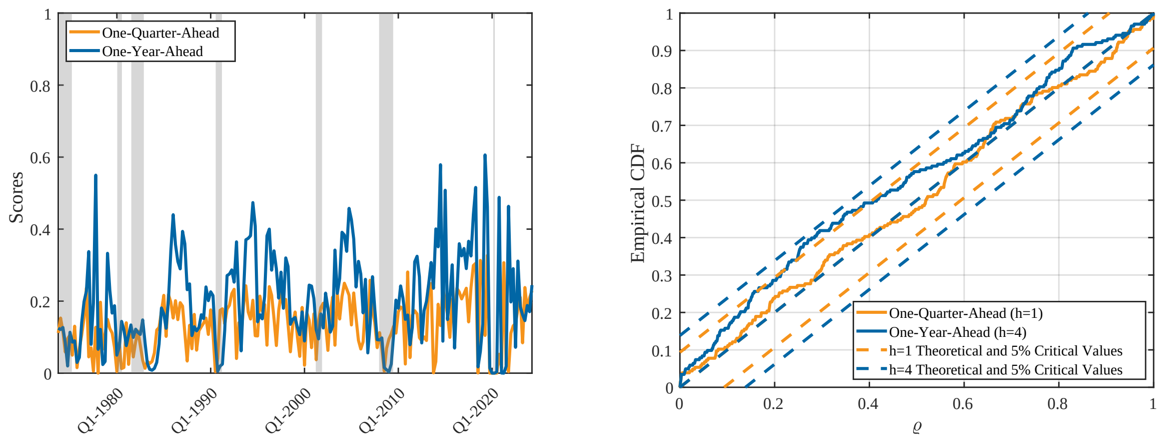
Panel (b) of Figure 4 shows the empirical cumulative distribution of the Probability Integral Transforms (PITs) to assess the calibration of the density forecasts. PIT measures the percentage of observations that are below any given quantile. The model is better calibrated the closer the empirical cumulative distribution of the PITs is to the 45-degree line. At the 45-degree line, the fraction of realizations below any given quantile $Q_{\varrho}(y_{t+h|t}|x_t)$ of the density forecasts is exactly equal to the quantile index value, ϱ. Following Rossi and Sekhposyan (2019), we report confidence bands around the 45- degree line to account for sample uncertainty. They test the null hypothesis of a dynamically correctly specified cumulative distribution function (CDF). This null hypothesis is not rejected in all our baseline density forecasts which means they are reliable from a calibration perspective.

5.2 Functional Regression Estimates

Table 1 reports our functional regression estimates for the *ur* specification using one-year-ahead GDP growth forecasts. These estimates are for inference purposes and use all our sam-

Figure 4: Density Forecast Out-Of-Sample Evaluation

(a) Predictive Scores (b) Probability Integral Transforms



Note: The predictive scores are the values obtained by evaluating the predictive distribution generated by a model at the realized value of the time series. The gray rectangles report recession dates as classified by the NBER. For calibration, we report the empirical cumulative distribution of the probability integral transform. Critical values are obtained as in Rossi and Sekhposyan (2019). It tests the null hypothesis of a dynamically correctly specified CDF.

ple from 1973Q1 to 2024Q3 with revised data. They include a 90% bootstrapped confidence interval computed using algorithm detailed in Appendix A. Results for one-quarter-ahead specifications are shown in Appendix C.

The estimates for $\hat{\alpha}$ suggest a point forecast, $m_{t+h|t}$, weight of 0.40 in our baseline unrestricted specification, *ur*. It means that we should not give full weight to the point forecast when the information of the density forecast is included, but the point forecast itself is still informative since its coefficient is statistically different from zero. The sign of this coefficient can be interpreted in the standard way, implying that increases in the point forecast are associated with increases in the one-year-ahead GDP growth rate, as it is well expected when the coefficient is positive. The magnitude of this coefficient should be interpreted as a weight, as it is usual in the forecast combination literature.

The OLS estimates for $f_{t+h|t}^{FPC_1}$ in specification *ur* tells us that the leading factor coefficient is statistically significant with a coefficient of 0.78. The factors are identified up to a sign meaning we cannot interpret the sign of coefficients of the factors alone; we need the shape of the associated factor and its loadings too, similar to our discussion in Figure 2. Thus, it is informative to recover the respective functional coefficient estimates, $\widehat{\beta}$, for interpretation purposes.

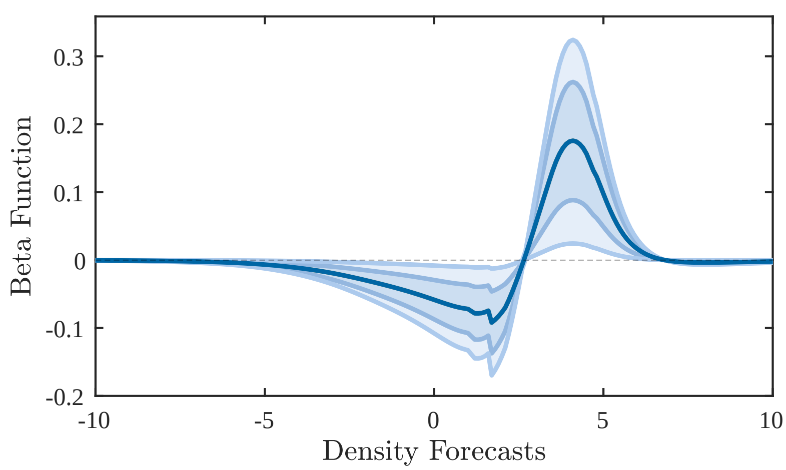
Only one factor was needed for the unrestricted specification. The number of factors was chosen with the BIC meaning less covariates fit better when estimating jointly the weights

Table 1: Functional Regression Estimates For One-Year-Ahead

	ur
$m_{t+h\|t}$	0.3970
	[0.2244, 0.5793]
$f_{t+h\|t}^{FPC_1}$	0.7779
	[0.1077, 1.4347]
R^2	0.2219
K	1
Observations	203

Note: The table shows the estimated OLS regression coefficients in the respective functional regression with sample 1973Q1-2024Q3. Covariates include an intercept, the point forecast, and the first K leading functional principal components of the density forecasts. The density forecasts are computed using linear quantile regressions of the one-year-ahead real GDP growth on current real GDP growth and NFCI; then, the resulting quantiles are used to fit a skewed t-distribution. We report 90% confidence intervals computed using 1,000 bootstrapped samples following the algorithm detailed in Appendix A.

Figure 5: Beta Function Estimates For One-Year-Ahead ur Model



Note: The figure shows the estimated functional coefficients in the respective functional regression with sample 1973Q1-2024Q3 and revised data. The density forecasts are computed using linear quantile regressions of the one-year-ahead real GDP growth on current real GDP growth and NFCI; then, the resulting quantiles are used to fit a skewed t-distribution. We report 68% and 90% confidence bands computed using 1,000 bootstrapped samples following the algorithm detailed in Appendix A.

for the point forecast and the first factor.

Figure 5 shows the functional coefficient estimates for the ur specification. The degree of simplicity of the functional coefficient depends on how many factors were used for estimation in the functional regression. Recall that the functional coefficient estimate can be recovered by using the inverse of the isometry as shown in equation 7. With K factors, each denoted v_k, and $\widehat{(\beta_k)}$ the respective OLS coefficient in the functional regression, recovering the functional coefficient is simply computing $\widehat{\beta} = \sum_{k=1}^{K} \widehat{(\beta_k)} \times v_k$. Hence, the domain of $\widehat{\beta}$ is the same as for each v_k, and the codomain of $\widehat{\beta}$ is the sum of the scaled codomains of each v_k with scale magnitude $\widehat{(\beta_k)}$.

Looking at Figure 5, the functional coefficient estimate for $\hat{\beta}$ in specification *ur* suggest that the effect of $\left(f_{t+h|t}\right)$ moves the distribution of forecasts to the right, by suppressing smaller forecasts less than about 2.5%, while increasing the occurrence of larger forecasts above 2.5%. This is a statistically significant shift in the point forecast suggested by the predictive content of the density forecast, which is associated with recession and expansion periods, as discussed before.

The specific magnitude of the shift depends on which value of the growth rate we are looking at. For example, a growth rate of around 3% should be shifted upwards and be around 3.2% instead. We can use the function as a way to correct the point forecast, with the corrected value incorporating the information contained in the respective factors of the density forecast. This could also be regarded as a way to align the point forecast with the information provided by the density forecast.

Results from the *ur* specification suggest the SPF as point forecast is conservative conditional on the extra information obtained from the density forecast. Point forecast below 2.5% should be adjusted downwards, whereas point forecast above 2.5% should be adjusted upwards.

6 Real-Time Forecasting Environment

The functional regression estimate results showed in the previous Section 5 were based on estimates using revised (last vintage) data. Nonetheless, a fair comparison between the point forecast, $m_{t+h|t}$, and our proposed method should use the information set available to survey respondents at each period. It coincides with a "real-time" forecasting exercise, i.e., an application in which the estimations are perform as if a forecaster is computing them every time new observations are available. These observations is what we call vintage data. We follow the standard practice in the forecasting literature of using the first vintage of the quarter to represent the quarter's information.

The number of factors used for the real-time forecasting exercise are chosen with BIC using the first forecasting and estimation window with data from 1973Q1 to 2004Q4. This number could be different from the results showed in the previous Section 5 for inference purposes for two reasons. First, the estimation sample before was 1971Q1 to 2024Q3 including the financial crisis and the Covid-19 pandemic. Second, the data used before was revised data, whereas the real-time forecasting exercise uses vintage data to avoid a forward-looking bias.

6.1 Out-of-Sample Forecasting Evaluation

Table 2 shows the results from our OOS forecasting evaluation for our functional regression specification with vintage data for one-year-ahead GDP growth forecasts. Appendix D shows results for one-quarter-ahead forecasts. We compute the bias, variance, and the root mean squared (forecasting) error (RMSE) relative to those of the point forecast, ($m_{t+h|t}$), in the denominator. It means that improvements are measured with values below one. We also report the Diebold-Mariano test statistic (DM) and its significance level for the null hypothesis of equal forecasting performance relative to the point forecast, $m_{t+h|t}$. A negative statistic value means the proposed regression has smaller forecasting errors relative to $m_{t+h|t}$. The estimation and forecasting procedure follows a rolling window approach as detailed in Section 4.

We also report results from two simple approaches using regression and OLS estimation. These regressions could be regarded as a direct forecasting procedure since we regress the target variable, y_{t+h}, on specific covariates. We want to address whether our method using

Table 2: Forecast Performance Comparison

	Relative to SPF (in denominator)				
	Bias	Var	RMSE	DM test	K
ur	0.2551	0.9657	0.6640	-3.8618***	2
Covariates	0.2702	1.5115	0.8205	-1.2138	0
Interquantile	0.3359	1.0329	0.7048	-3.7011***	0

Note: The table shows out-of-sample forecast performance measures of the forecast (fitted) values from the functional regressions relative to the point forecast (SPF), $m_{t+h|t}$, for the target variable and forecasting horizon one-year-ahead ($h = 4$). The out-of-sample is 2004Q1-2024Q3. The functional regressions estimation sample is 1973Q1-2003Q4 for GDP growth in the first rolling window. ur means unrestricted functional regression specification. The bias, variance, and root mean squared (forecasting) error (RMSE) are reported in proportion to those of the point forecast (SPF), $m_{t+h|t}$, which is included in the denominator. DM test is the Diebold-Mariano test statistic with the point forecast (SPF), $m_{t+h|t}$, as benchmark. ***, **, and * mean the respective forecast (fitted) value from the functional regression specification rejects the null hypothesis of equivalent performance at the 1%, 5%, and 10% significance level, respectively. A negative (positive) DM number means smaller (larger) errors relative to $m_{t+h|t}$. K is the number of functional principal components used in the respective functional regression as an approximation of the density forecasts. K was chosen with the BIC using the sample in the first rolling window.

density forecast improves also relative to simple regression alternatives. First, we use the point forecast together with the covariates used for density forecasting, i.e. values of real GDP growth rate and the NFCI, but here we use the covariates directly when estimating a simple regression for the target variable. Second, we compute the inter-quartile range, i.e. the 75-percentile value minus the 25-percentile value, from the density forecasts, and use it as a covariate together with the point forecast to compute a direct forecast with simple regression for the target variable. These alternatives also use vintage data and a rolling window forecasting scheme in the same fashion as in our baseline specification.

Table 2 shows that the improvement of point forecasts using density forecasts is possible. Our proposed functional regression specification have less bias with only 26% the bias of $m_{t+h|t}$. The variance is smaller by 3% only. These result in a decrease of the RMSE of around 33% relative to $m_{t+h|t}$ for our *ur* specification. The associated DM test statistic being negative and with significance level below 1% shows that the improvement is statistically significant. Using the covariates directly and jointly with the point forecast does not provide improvement relative to using the point forecast alone. In contrast, using the inter-quartile range from the density forecast together with the point forecast provides improvement of near 29% in the RMSE due to a decrease in the bias. This confirms that the predictive content of the density forecast for one-year-ahead forecasting horizon comes from the information beyond the point forecast, $m_{t+h|t}$.

Our results suggest that the best way to improve the SPF's (consensus) average point forecast for one-year-ahead GDP growth is by using it jointly with density forecast. It seems that the improvement comes from exploiting the predictive content from the density forecast related to recession and expansion periods (see Figure 2). The functional regression coefficient estimate from the *ur* specification, shown in Figure 5, suggests that the SPF's consensus (average) forecast tends to be conservative, i.e., low growth rate forecasts should be adjusted downwards, whereas high growth rate forecasts should be adjusted upwards. This conservativeness is prominent for recessions and expansion periods because it is difficult to forecast those extreme events as they are far away from the expected value, i.e. the center of the distribution. Exploiting the information content from the tail of the distribution is the key aspect in the growth-at-risk literature and others similar. It helps the forecasting performance in our case.

7 Robustness Checks

We perform robustness checks as well as additional estimation, inference, and real-time forecasting exercises. These are related to studying how the density forecast provides improvement by using variations of our functional regression, forecasting the one-quarter-ahead GDP growth rate, changing the starting point for the rolling window forecasting scheme, and checking whether the estimated functional coefficients change when using different ways of computing density forecasts.

We first start explaining about different functional regression specifications. We consider giving full weight to the point forecasts, $\left(m_{t+h|t}\right)$, by estimating a *restricted functional regression*, or rr for short, of the form

$$y_{t+h} = \tau + m_{t+h|t} + \left\langle \beta, f_{t+h|t} \right\rangle + \varepsilon_{t+h} \qquad (rr)$$

where we assume $\alpha = 1$. Here, we want to measure whether $\left(f_{t+h|t}\right)$ can further explain systematic information unexplained by $\left(m_{t+h|t}\right)$. The specification in equation (rr) is explicitly assuming that the point forecast, $\left(m_{t+h|t}\right)$, is our main prediction and it is correct on average. If the point forecast is a good estimator of the expected value of the time series for the target variable, then, a full weight is appropriate and the density forecast will suggest adjustments only for realizations far away from the expected value, usually, realized values in the tails of the distribution.

It is also possible to further investigate the way density forecast improves the point forecast by separating $\left(f_{t+h|t}\right)$ into what we call a *mean factor*, $\left(\mu\left(f_{t+h|t}\right)\right)$, and *centered* density, $\left(f_{c,t+h|t}\right)$, to forecast $\left(y_{t+h|t}\right)$.[16] This decomposition gives us the predictive content of $\left(\mu\left(f_{t+h|t}\right)\right)$ and $\left(f_{c,t+h|t}\right)$ to predict the target variable beyond the point forecast.

To do that, we set the first factor of the density forecast $\left(f_{t+h|t}\right)$ to be $u_m = r$ meaning that $\left\langle u_m, f_{t+h|t} \right\rangle = \int r f_{t+h|t}(r)dr = \mu\left(f_{t+h|t}\right)$ and we call u_m the *mean factor*.[17] Then, we define a *centered* density forecast $\left(f_{c,t+h|t}\right)$ as $f_{c,t+h|t} = f_{t+h|t}\left(r + \mu\left(f_{t+h|t}\right)\right)$.

Therefore, we also estimate an unrestricted specification of the form

$$y_{t+h} = \tau + \alpha\, m_{t+h|t} + \beta_m\, \mu\left(f_{t+h|t}\right) + \left\langle \beta_c, f_{c,t+h|t} \right\rangle + \varepsilon_{t+h} \qquad (u\mu r)$$

which we called *unrestricted with mean factor functional regression* or $u\mu r$ for short. Here, $\left(\mu\left(f_{t+h|t}\right)\right)$, $\left(f_{c,t+h|t}\right)$, and $\left(m_{t+h|t}\right)$ are jointly used to forecast $\left(y_{t+h}\right)$. Like the previous unrestricted specification, this allows the jointly use of the centered density forecast, the

[16] Here, we impose the mean factor as the first functional principal component. Therefore, in the Hilbert space, before using the isometry, this functional is indeed a mean factor being a function and forming a different basis together with the other factors from the centered density. It is not the most effective representation of the temporal variation of the functional covariate in functional regressions, however, as stated in Subsection 2.2.

[17] To be precise, u_m is deliberately inserted as the first basis function (first factor), and it is associated with the loadings represented by $\mu(f_{t+h|t})$. These loadings are then included in the regression alongside the empirical FPC loadings derived from the centered densities. We call $\mu(f_{t+h|t})$ the *mean factor* for simplicity.

mean factor, and the point forecast by giving agnostic weights.

Additionally, we consider the restricted version called the *restricted with mean factor functional regression*, or $r\mu r$ for short, which is the following equation

$$y_{t+h} = \tau + m_{t+h|t} + \beta_m \mu\left(f_{t+h|t}\right) + \left\langle \beta_c, f_{c,t+h|t} \right\rangle + \varepsilon_{t+h} \qquad (r\mu r)$$

where $\alpha = 1$. Similar to the previous restricted specification, this specification permits the centered density forecast and the mean factor to take only a complimentary role in predicting the target variable relative to the point forecast.

It is feasible to obtain estimates, perform inference, and real-time forecasting exercises for each of the functional regression equations (ur), (rr), ($u\mu r$), and ($r\mu r$), by using the estimation approach detailed in Subsection 2.2, the methodology outlined in Section 4, and performing the bootstrap procedure outlined in Appendix A. The only difference is in equations ($u\mu r$) and ($r\mu r$) where we impose the mean factor, $\left(\mu\left(f_{t+h|t}\right)\right)$, as the first functional principal component, with the rest of the FPCs being those obtained from the sample variance operator, $\bar{\Gamma}$, of the centered densities with the procedure explained previously. For these specifications, the coefficient β_m and its covariate $\mu\left(f_{t+h|t}\right)$ are standard, i.e., $\beta_m \in \mathbb{R}$ and $\mu\left(f_{t+h|t}\right) \in \mathbb{R}^{T-h}$.

While specifications from equation ($u\mu r$) and equation ($r\mu r$) gives us a way to measure the predictive content of the mean factor of the density forecast, $\left(\mu\left(f_{t+h|t}\right)\right)$, beyond that of the point forecast, $\left(m_{t+h|t}\right)$, it is possible that $\left(\mu\left(f_{t+h|t}\right)\right)$ and $\left(m_{t+h|t}\right)$ are highly correlated. If that is the case, the resulting collinearity would increase the variance of the estimated coefficients and of the resulting point forecast, $\left(\hat{y}_{t+h|t}\right)$. Then, $\left(\mu\left(f_{t+h|t}\right)\right)$ and $\left(m_{t+h|t}\right)$ would provide similar predictive content and only one of them should be used. Furthermore, other specifications could be preferred by the out-of-sample forecasting performance evaluation meaning that it would be better to use the density forecast without separating the mean factor.

Estimation and inference results for all our proposed functional regression specifications as well as for both forecasting horizons (one-quarter-ahead, $h = 1$, and one-year-ahead, $h = 4$) are shown in Appendix C. For functional regression estimates with one-year-ahead GDP growth rates, it is better to use jointly the point and density forecast in an unrestricted fashion without separating the mean factor. The mean factor is highly correlated with the point forecast leading to increased estimation variance when separating the mean factor and using it together with the centered density (see Table A1).

Furthermore, imposing the restriction of full weight to the point forecast results in the functional coefficient not being statistically significant. In contrast, the density forecast provides information useful related to recession and expansion periods in the unrestricted, ur, specification (see Figure A1).

For one-quarter-ahead specifications, it seems that the point forecast has more predictive content as it has higher weight estimates. They do not reject the null hypothesis of unit weight in the unrestricted specifications (see Table A2). The predictive content of the density forecast here is concentrated in the mean factor, but as it is highly correlated with the point forecast, there is overlapping predictive content that increases the forecasting error variance. Here, all the functional coefficient are not statistically significant (see Figure A2).

The respective real-time forecasting performance evaluation results are shown in Appendix D. The unrestricted specification, ur, is the best performer for one-year-ahead forecasts mainly due to a decrease in forecasting bias (see Table A3). There is no improvement for one-quarter-ahead forecast and only our restricted specifications have similar forecasting performance as using the point forecast alone (see Table A3). This is expected given the overlapping predictive content between the point forecast and the mean factor for one-quarter-ahead, whereas for one-year-ahead, it is the information related to recessions and expansions contained in the tails of the density forecast what helps in the improvement.

Other robustness checks are shown in Appendix E or in the Supplementary Material. First, in the Appendix E we consider one subsample case. We use estimation samples from 1973Q1 to 1994Q4 or 1973Q1 to 1994Q1 for $h = 1$ and $h = 4$, respectively, with first OOS date 1995Q1. In other words, we estimate our functional regressions with in-sample data until 1994Q4 in the first rolling window. The OOS sizes change accordingly with end date 2024Q3. It gives us a total of 123 OOS observations. It means that the OOS size is bigger for computing forecasting evaluation measures, but at the cost of including less information for obtaining coefficient estimates.

Table A4 shows results similar to those shown in our baseline rolling window procedure (when the first OOS observation in the rolling window starts at 2004Q1). There is less bias relative to the point forecast, $m_{t+h|t}$, for one-year-ahead GDP growth forecast in all our functional regression specifications. The variance is also smaller for our ur specification (see Panel (b) in Table A4) being the best specification in terms of forecasting accuracy with a RMSE of around 65% relative to the point forecast. We don't see any improvement for one-quarter-ahead forecasts.

The Supplementary Materials include several parametric and nonparametric ways to compute density forecast with and without a data-rich environment. Hence, we have different versions of our proposed functional regressions. All the specifications we consider there are with revised data due to data limitations, meaning they suffer from a forward-looking bias. We also include there results of our baseline specifications (few linear quantiles using the

GDP growth rate and NFCI) showed here, but with revised data for comparison purposes.

Interestingly, we find that data-rich machine learning methods such as quantile Lasso regression and Generalized Random Forests, although appealing in their ability to handle many covariates, do not outperform simpler linear quantile regressions with carefully chosen predictors. The forecasting performance results are similar. This suggests that in macroeconomic forecasting, where sample sizes are relatively limited and structural breaks may be present, parsimony and interpretability remain crucial. In particular, the gains from flexible methods may be offset by the noise introduced through large covariate spaces when estimating predictive distributions.

In the Supplementary Material, we also check whether our coefficient estimates from our functional regressions have similar magnitude and sign when we use the different ways of obtaining the density forecast. We focus only in the one-year-ahead GDP growth forecasting *ur* specification because our results suggest statistically significant improvements with it. The coefficient estimate $\hat{\alpha}$ for the $m_{t+h|t}$ weight is robust being around 0.39. Also, the shape of the functional coefficient estimate, $\hat{\beta}$, is robust. As previously discussed, the functional coefficient suppresses smaller forecasts less than 2.5%, while increases the occurrence of larger forecasts above 2.5%. This predictive content of the density forecast is associated with recession and expansion periods in all cases.

8 Conclusions

Our findings suggest that it is possible to improve a given point forecast with density forecasts with our proposed functional regression approach. However, the improvement depends on the forecasting horizon. In our applications to U.S. quarterly GDP growth, it offers substantial gains in one-year-ahead forecast accuracy when using the SPF consensus as point forecast, which is widely used in practice. There is not improvement for one-quarter-ahead forecasting.

The use of functional principal components as a basis enables the method to extract the relevant predictive content of the density forecast in an effective way. Our results indicate that the predictive content of densities varies across the forecasting horizon: for one-year-ahead GDP growth, capturing the asymmetry associated with business cycles significantly reduces forecast bias, while for one-quarter-ahead, the central mass of the density–what we define as the "mean factor"–primarily drives the predictive content from the density forecast.

Future research could explore hybrid approaches that combine the strengths of machine learning with the interpretability of functional regression models. For instance, variable selection or flexible nonparametric methods tailored to density forecasting with potential sparsity or adaptive shrinkage techniques could enhance the robustness of the model in high-dimensional settings. These methods applied directly to functional regression could also be useful when including several different point or density forecasts.

Additionally, applying the methodology to other macro-financial variables such as inflation, unemployment, interest rates, credit spreads, or labor market indicators could shed further light on the role of distributional information, forecasts, and economic expectations. Incorporating real-time data updates and evaluating the performance of the method in nowcasting environments would also be valuable extensions for practical policy applications.

References

Adrian, T., Boyarchenko, N., and Giannone, D. (2019). Vulnerable growth. *American Economic Review*, 109(4):1263–1289.

Adrian, T., Grinberg, F., Liang, N., Malik, S., and Yu, J. (2022). The term structure of growth-at-risk. *American Economic Journal: Macroeconomics*, 14(3):283–323.

Amburgey, A. and McCracken, M. W. (2023a). *Growth-at-risk is Investment-at-risk*. Federal Reserve Bank of St. Louis, Research Division.

Amburgey, A. J. and McCracken, M. W. (2023b). On the real-time predictive content of financial condition indices for growth. *Journal of Applied Econometrics*, 38(2):137–163.

Ang, A., Bekaert, G., and Wei, M. (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of monetary Economics*, 54(4):1163–1212.

Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65(2):367–389.

Blinder, A. S. (1999). *Central banking in theory and practice*. Mit press.

Bosq, D. (2000). *Linear processes in function spaces: theory and applications*, volume 149. Springer Science & Business Media.

Carriero, A., Clark, T. E., and Marcellino, M. (2024a). Capturing macro-economic tail risks with bayesian vector autoregressions. *Journal of Money, Credit and Banking*, 56(5):1099–1127.

Carriero, A., Clark, T. E., and Marcellino, M. (2024b). Specification choices in quantile regression for empirical macroeconomics. *Journal of Applied Econometrics*.

Chang, M., Chen, X., and Schorfheide, F. (2024). Heterogeneity and aggregate fluctuations. *Journal of Political Economy*, 132(12):4021–4067.

Chang, Y., Durlauf, S. N., Lee, S., and Park, J. Y. (2023). A trajectories-based approach to measuring intergenerational mobility. Technical report, National Bureau of Economic Research.

Chang, Y., Kim, C. S., and Park, J. Y. (2016). Nonstationarity in time series of state densities. *Journal of Econometrics*, 192(1):152–167.

Chang, Y., Park, J. Y., and Pyun, D. (2021). From functional autoregressions to vector autoregressions. Technical report, Working Paper, Indiana University.

Clark, T. and McCracken, M. (2013). Advances in forecast evaluation. *Handbook of economic forecasting*, 2:1107–1201.

Clark, T. E. and McCracken, M. W. (2012). Reality checks and comparisons of nested predictive models. *Journal of Business & Economic Statistics*, 30(1):53–66.

Corradi, V., Fosten, J., and Gutknecht, D. (2023). Out-of-sample tests for conditional quantile coverage an application to growth-at-risk. *Journal of econometrics*, 236(2):105490.

Corradi, V. and Swanson, N. R. (2006). Predictive density evaluation. *Handbook of economic forecasting*, 1:197–284.

Croushore, D. (2010). An evaluation of inflation forecasts from surveys using real-time data. *The BE Journal of Macroeconomics*, 10(1).

Croushore, D., Stark, T., et al. (2019). Fifty years of the survey of professional forecasters. *Economic Insights*, 4(4):1–11.

Diebold, F. X. and Mariano, R. S. (1995). Com paring predictive accu racy. *Journal of Business and Economic Statistics*, 13(3):253–263.

Faust, J. and Wright, J. H. (2013). Forecasting inflation. In *Handbook of economic forecasting*, volume 2, pages 2–56. Elsevier.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, 79(2):453–497.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50.

Lopez-Salido, D. and Loria, F. (2024). Inflation at risk. *Journal of Monetary Economics*, 145:103570.

Rossi, B. and Sekhposyan, T. (2019). Alternative tests for correct specification of conditional predictive densities. *Journal of Econometrics*, 208(2):638–657.

Stark, T. et al. (2010). Realistic evaluation of real-time forecasts in the survey of professional forecasters. *Federal Reserve Bank of Philadelphia Research Rap, Special Report*, 1.

West, K. D. (2006). Forecast evaluation. *Handbook of economic forecasting*, 1:99–134.

Appendix A Bootstrap For Functional Regressions

The regression in equation (6) is a standard regression. We can use the standard wild bootstrap procedure for regressions to obtain uncertainty measurements and for inference purposes. Specifically, we construct the artificial samples and bootstrap critical values using a modification of the bootstrap procedure for functional regressions proposed by Chang et al. (2021) following Clark and McCracken (2012).[18]

The steps for the bootstrap procedure are the following:

Step 1. Using OLS, estimate the parameters and store the residuals $\hat{\varepsilon}_{t+h}$ for equation 6. Compute and store the coefficients from the equivalent equation

$$y_{t+h} \approx \tau + \alpha m_{t+h|t} + \langle \beta \rangle' \langle f_{t+h|t} \rangle + \varepsilon_{t+h}$$

with $\tau = \bar{y} - \alpha \bar{m} - \langle \beta \rangle' \langle \bar{f} \rangle$.

Step 2. If $h \geqslant 2$, use Nonlinear Least Squares (NLLS) to estimate an $WA(h-1)$ model for the OLS residuals $\hat{\varepsilon}_{t+h}$ such that $\varepsilon_{t+h} = \epsilon_{t+h} + \theta_1 \epsilon_{t+h-1} + \ldots + \theta_{h-1} \epsilon_{t+1}$.

Step 3. Let η_{t+h}, $t = 1, \ldots, T-h$, denote an i.i.d $N(0,1)$ sequence of simulated random variables. If $h = 1$, define $\hat{\varepsilon}^*_{t+1} = \eta_{t+1} \hat{\varepsilon}_{t+1}$, $t = 1, \ldots, T-1$. If $h \geqslant 2$, define $\hat{\varepsilon}^*_{t+h} = \left(\eta_{t+h} \hat{\epsilon}_{t+h} + \hat{\theta}_1 \eta_{t-1+h} \hat{\epsilon}_{t+h-1} + \ldots + \hat{\theta}_{h-1} \eta_{t+1} \hat{\epsilon}_{t+1} \right)$, $t = 1, \ldots, T-h$.

Step 4. Form artificial samples of y^*_{t+h} using the fixed regressor structure,

$$y^*_{t+h} = \hat{\tau} + \hat{\alpha} m_{t+h|t} + \langle \hat{\beta} \rangle' \langle f_{t+h|t} \rangle + \hat{\varepsilon}^*_{t+h}.$$

Step 5. Repeat steps 3 - 4 a large number of times: $j = 1, \ldots, N$. We define $N = 1,000$.

[18]Note that the same factor $\eta_{t+h-j} \hat{\epsilon}_{t+h-j}$ enters several adjacent bootstrap residuals in Step 3 of our bootstrap procedure. This overlap ensures that the simulated series retains the correct first-through-$(h-1)$-order autocovariances, reproducing the moving average structure used in Step 2. For more details see Clark and McCracken (2012).

Appendix B Point Forecasts

We use the (consensus) mean value of the responses from the Survey of Professional Forecasters (SPF) as point forecast. The SPF is the oldest quarterly survey of macroeconomic forecasts in the United States. It has been performed since 1968 and the Federal Reserve Bank of Philadelphia has conducted it since the second quarter of 1990.

The survey polls professional economists each quarter on their views about the economy over the next few years. The survey participants cover a wide spectrum of industries and it provides forecasts for nominal GDP and the GDP price index, corporate profits, real GDP and its components, and a number of monthly business indicators, such as interest rates, housing starts, industrial production, and the consumer price index.[19] We focus on the real GDP.

Stark et al. (2010) study the performance of the SPF. Overall, the survey accuracy falls when the forecasting horizon increases, data revisions can have a large effect on the survey accuracy, and the survey projections generally outperform univariate autoregressive time-series models at short horizons. In general, there is a consensus that SPF point forecasts are difficult to beat in real-time accuracy (Ang et al., 2007; Croushore, 2010; Faust and Wright, 2013; Croushore et al., 2019).

The SPF has maintained a steady schedule since the third-quarter survey of 1990 to align with the Bureau of Economic Analysis (BEA)'s advance release of the data.[20] The BEA's report is issued each late January, April, July, and October, and includes the first estimate of the historical realization of variables from the national accounts. Survey respondents bring their projections around the middle of each quarter, before the BEA releases, and one month later the first revision of the previous quarter advance estimate. It is reasonable to assume the survey participants know the values of interest rates and labor-market variables for the first month of each quarter, but not for housing starts, industrial production, and the consumer price index when they form their projections (Stark et al., 2010).

In each quarter, the SPF asks for the expected Real GDP level directly for the current quarter (nowcasting) and the next four quarters –a forecasting horizon going from one-quarter-ahead up to one-year-ahead. It also reports the mean value by computing the average

[19]Respondents belong to Insurance, Investment Banking, Commercial Banking, Payment Services, Hedge Funds, Mutual Funds, Association of Financial Service Providers, Asset Management, Manufacture, Universities, Forecasting Firms, Investment Advisors, Pure Research Firms, Consulting Firms, among other industries.

[20]Previously, the SPF was conducted by the National Bureau of Economic Research (NBER) and the American Statistical Association (ASA). As mentioned by Stark et al. (2010), the survey timing was less certain prior to the second quarter of 1990. Nonetheless, internal analysis by Philadelphia Fed's Real-Time Data Research Center suggests that the NBER/ASA's schedule was about the same as the current SPF schedule.

from individual responses for each forecasting horizon. We use these mean values.

For real GDP growth, let h be the forecast horizon for one-quarter-ahead ($h = 1$) or one-year-ahead ($h = 4$), and let t the period of the survey. For one-year-ahead, we compute a survey-based year-over-year expected real GDP growth rate. For a one-quarter-ahead, we compute a survey-based annualized quarter-over-quarter real GDP growth rate. In both, we use the mean values of the real GDP level obtained from the survey performed at time t as follows:

$$m^g_{t+h|t} \equiv \hat{g}_{t+h|t-1} = 100 \times \left[\left(\frac{\hat{y}_{t+h|t-1}}{\hat{y}_{t|t-1}} \right)^{5-h} - 1 \right], \;\; h \in \{1, 4\}$$

where $m^g_{t+h|t}$ represent the point forecast for the expected value of the real GDP growth rate that will be observed at time $t + h$ conditional on the information known in period t. $\hat{y}_{t+h|t-1}$ is the expected real GDP level for period $t+h$ (one-quarter-ahead forecast for $h = 1$ or one-year-ahead forecast for $h = 4$) and $\hat{y}_{t|t-1}$ is the expected real GDP level for period t (nowcast), respectively.[21] We will refer to point forecasts for GDP growth at horizon h, $m^g_{t+h|t}$, only as $m_{t+h|t}$ for simplicity.

[21]We considered other alternatives to the SPF for survey-based growth rates. However, other sources like Federal Open Market Committee (FOMC) announcements and their growth forecasts are not available at the quarterly frequency or the sample is too small. Usually, they provide growth forecast for the end of the year, not quarter-over-quarter or year-over-year growth forecast.

Appendix C Estimation and Inference Results

Tables A1 and A2 report our functional regression estimates for ur, rr, $u\mu r$, and $r\mu r$ specifications for the one-year-ahead and one-quarter-ahead GDP growth forecasts, respectively. These estimates are for inference purposes and use all our sample from 1973Q1 to 2024Q3 with revised data. They include a 90% bootstrapped confidence interval computed using the algorithm detailed in Appendix A.

For one-year-ahead, the estimates reported in Table A1 for $\hat{\alpha}$ suggest point forecast, $m_{t+h|t}$, weights of 0.40 and 0.32 in our unrestricted specifications, ur and $u\mu r$, respectively. It means that we should not give full weight to the point forecast when the information of the density forecast is included, but the point forecast itself is still informative since its coefficient is statistically different from zero.

The mean factor, $\left(\mu\left(f_{t+h|t}\right)\right)$, has a positive and statistically significant coefficient of 0.81 in specification $u\mu r$. However, the coefficient estimate for specification $r\mu r$ is -0.16 and not statistically different from zero. The sign of these coefficients can be interpreted in the standard way, implying that increases in the mean factor are associated with increases in the one-year-ahead GDP growth rate, as it is well expected when the coefficient is positive. The magnitude of these coefficients should be interpreted as a weight, as it is usual in the forecast combination literature.

The $u\mu r$ specification tell us that both the point forecast and the mean factor are informative, but none should be given full weight. On the contrary, specification $r\mu r$ tell us that, once we have given full weight to the point forecast, the mean factor does not provide additional predictive content. These are not contradictory results since a joint estimate using both the point forecast and the mean factor is different from a restriction placed on the point forecast coefficient. The mean factor is highly correlated with the point forecast, with simple correlation coefficient of 0.59. This leads to an increase in the variance estimates from specification $u\mu r$. It is also consistent with giving full weight to the point forecast and then the mean factor not providing additional predictive content. Our forecast evaluation results shown below support this view since the $u\mu r$ specification has higher forecasting error variance than the $r\mu r$ specification. Therefore, the density forecast would improve the forecasting accuracy for one-year-ahead forecasts if it provides information beyond the mean factor.

The OLS estimates for $f_{t+h|t}^{FPC_1}$ in specifications ur and rr tell us that the leading factor coefficient is statistically significant. Specification ur shows a coefficient of 0.78 and specification rr a coefficient of -0.71.

The factors are identified up to a sign meaning we cannot interpret the sign of coefficients

Table A1: Functional Regression Estimates For One-Year-Ahead GDP Growth Forecasts

	ur	rr	$u\mu r$	$r\mu r$
$m_{t+h\|t}$	0.3970		0.3229	
	[0.2244, 0.5793]		[0.1629, 0.4738]	
$\mu(f_{t+h\|t})$			0.8052	-0.1570
			[0.2706, 1.4070]	[-1.0784, 0.8265]
$f_{t+h\|t}^{FPC_1}$	0.7779	-0.7052		
	[0.1077, 1.4347]	[-1.3676, -0.0537]		
$f_{t+h\|t}^{FPC_2}$		-0.9402		
		[-1.7028, -0.1522]		
$f_{t+h\|t}^{FPC_3}$		1.0630		
		[-0.3656, 2.5533]		
$f_{t+h\|t}^{FPC_4}$		0.5839		
		[-1.3304, 2.4886]		
$f_{c,t+h\|t}^{FPC_1}$			-0.4940	-0.4323
			[-1.5585, 0.4915]	[-2.1733, 1.2227]
$f_{c,t+h\|t}^{FPC_2}$				-0.8766
				[-1.6436, -0.0999]
$f_{c,t+h\|t}^{FPC_3}$				1.4433
				[-0.9872, 3.7375]
R^2	0.2219	0.0495	0.2522	0.0487
K	1	4	1	3
Observations	203	203	203	203

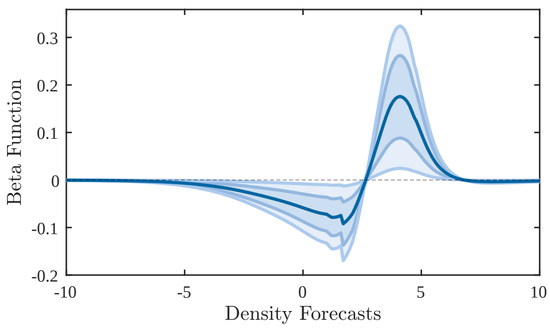
Note: The table shows the estimated OLS regression coefficients in the respective functional regression with sample 1973Q1-2024Q3. Covariates include an intercept, the point forecast, and the first K leading functional principal components or the mean factor, respectively, of the density forecasts. The density forecasts are computed using linear quantile regressions of the one-year-ahead real GDP growth on current real GDP growth and NFCI; then, the resulting quantiles are used to fit a skewed t-distribution. We report 90% confidence intervals computed using 1,000 bootstrapped samples following the algorithm detailed in Appendix A.

of the factors alone; we need the shape of the associated factor and its loadings too, similar to our discussion in Figure 2. Thus, it is informative to recover the respective functional coefficient estimates $\widehat{\beta}$, and $\widehat{\beta}_c$ for interpretation purposes.

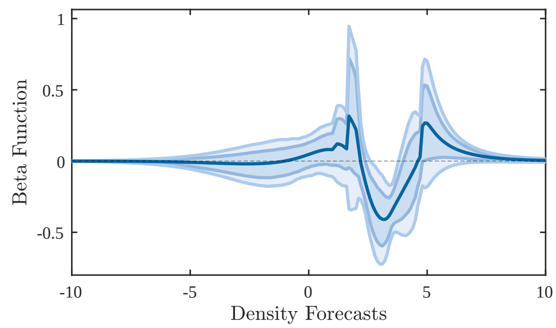
For specifications $u\mu r$ and $r\mu r$, the respective OLS estimates for $f_{c,t+h|t}^{FPC_1}$ are not statistically significant. It means they do not provide additional predictive content once we have controlled for the other factors, the point forecast, or the mean factor, respectively in these specifications.

Figure A1: Beta Function Estimates For One-Year-Ahead GDP Growth Forecasts

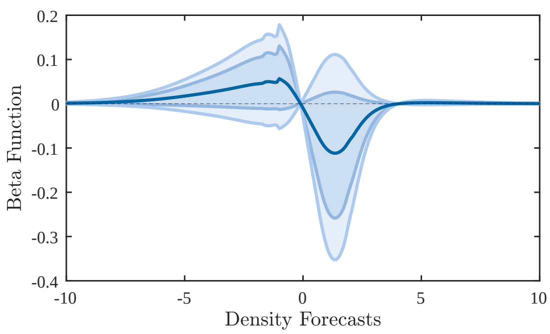
(a) Unrestricted Model



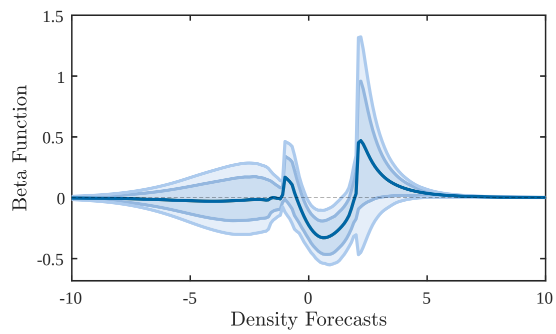
(b) Restricted Model



(c) Unrestricted Mean Factor Model



(d) Restricted Mean Factor Model



Note: The figure shows the estimated functional coefficients in the respective functional regression with sample 1973Q1-2024Q3 and revised data. The density forecasts are computed using linear quantile regressions of the one-year-ahead real GDP growth on current real GDP growth and NFCI; then, the resulting quantiles are used to fit a skewed t-distribution. We report 68% and 90% confidence bands computed using 1,000 bootstrapped samples following the algorithm detailed in Appendix A.

Only one factor was needed for unrestricted specifications, but more factors were added when imposing full weight on the point forecast in both of our restricted specifications. The number of factors was chosen with the BIC meaning less covariates fit better when estimating jointly the weights for the point forecast and the first factor. In contrast, the restriction of full weight to the point forecast implies that more information should be extracted from the density forecast, i.e. we need to include more factors, in order to improve fit.

Figure A1 shows the functional coefficient estimates for each specification. Some functions have simpler shapes, like that of Panel (a) for the *ur* specification, whereas others are more complex. The degree of simplicity depends on how many factors were used for estimation of the functional regression. Recall that the functional coefficient estimate can be recovered by using the inverse of the isometry as shown in equation 7. With K factors, each denoted v_k, and $\widehat{(\beta_k)}$ the respective OLS coefficient in the functional regression, recovering the functional coefficient is simply computing $\widehat{\beta} = \sum_{k=1}^{K} \widehat{(\beta_k)} \times v_k$. Hence, the domain of $\widehat{\beta}$ is the same as for each v_k, and the codomain of $\widehat{\beta}$ is the sum of the scaled codomains of each v_k with scale magnitude $\widehat{(\beta_k)}$.

Looking at Panel (a) in Figure A1, the functional coefficient estimate for $\hat{\beta}$ in specification ur suggest that the effect of $\left(f_{t+h|t}\right)$ moves the distribution of forecasts to the right, by suppressing smaller forecasts less than about 2.5%, while increasing the occurrence of larger forecasts above 2.5%. This is an statistically significant shift in the point forecast suggested by the predictive content of the density forecast, which is associated with recession and expansion periods, as discussed before.

The specific magnitude of the shift depends on which value of the growth rate we are looking at. For example, a growth rate of around 3% should be shifted upwards and be around 3.2% instead. We can use the function as a way to correct the point forecast, with the corrected value incorporating the information contained in the respective factors of the density forecast. This could also be regarded as a way to align the point forecast with the information provided by the density forecast.

Results from the ur specification suggest the SPF as point forecast is conservative conditional on the extra information obtained from the density forecast. Point forecast below 2.5% should be adjusted downwards, whereas point forecast above 2.5% should be adjusted upwards.

The functional coefficient from the $u\mu r$ specification is not statistically significant (see Panel (c) of Figure A1), as we expect from the OLS estimation results since the OLS estimate is not statistically significant in Table A1 for the leading factor of the centered density. Estimates for ur show the first factor contains predictive content, but specification $u\mu r$ shows the centered density does not play a role when separating the mean factor. These are not contradictory results for two reasons. First, the mean factor and the leading factor from the density forecast are not the same. Using the mean factor is not the most effective representation of the temporal variation as described in Section 4. Second, as it is expected, the mean factor is highly correlated with the leading factor of the centered density with simple correlation coefficient 0.84. It leads to collinearity problems when estimating $u\mu r$. The high correlation comes from the fact that the average is a measure of location, then, when the leading factor representing changes of the density forecast moves, the location or average also changes. This increases the variance of the estimated coefficient of the first factor of the centered density.

The functional coefficient estimates shown in Panels (b), for the rr specification, and (d) for the $r\mu r$ specification, have a more complex shape with some region in the middle

Table A2: Functional Regression Estimates For One-Quarter-Ahead GDP Growth Forecasts

	ur	rr	$u\mu r$	$r\mu r$
$m_{t+h\|t}$	0.9411		0.6292	
	[0.1144, 1.7616]		[0.0188, 1.2481]	
$\mu(f_{t+h\|t})$			3.0785	2.6779
			[0.8309, 5.5822]	[0.3131, 5.2122]
$f_{t+h\|t}^{FPC_1}$	-1.7204	-1.5452		
	[-5.8675, 2.2263]	[-3.9096, 0.5597]		
$f_{t+h\|t}^{FPC_2}$	2.0118	2.0045		
	[-0.2783, 4.4747]	[-0.3559, 4.5191]		
$f_{t+h\|t}^{FPC_3}$	0.6049			
	[-5.4910, 6.6635]			
$f_{c,t+h\|t}^{FPC_1}$			10.0557	9.5054
			[0.0012, 20.4076]	[-0.0597, 19.2240]
$f_{c,t+h\|t}^{FPC_2}$			1.9758	1.9399
			[-0.2881, 4.3974]	[-0.2545, 4.3287]
$f_{c,t+h\|t}^{FPC_3}$			-12.4177	-11.4135
			[-27.5601, 1.1296]	[-26.5007, 2.2094]
R^2	0.2108	0.0326	0.2832	0.1081
K	3	2	3	3
Observations	206	206	206	206

Note: The table shows the estimated OLS regression coefficients in the respective functional regression with sample 1973Q1-2024Q3. Covariates include an intercept, the point forecast, and the first K leading functional principal components or the mean factor, respectively, of the density forecasts. The density forecasts are computed using linear quantile regressions of the one-quarter-ahead real GDP growth on current real GDP growth and NFCI; then, the resulting quantiles are used to fit a skewed t-distribution. We report 90% confidence intervals computed using 1,000 bootstrapped samples following the algorithm detailed in Appendix A.

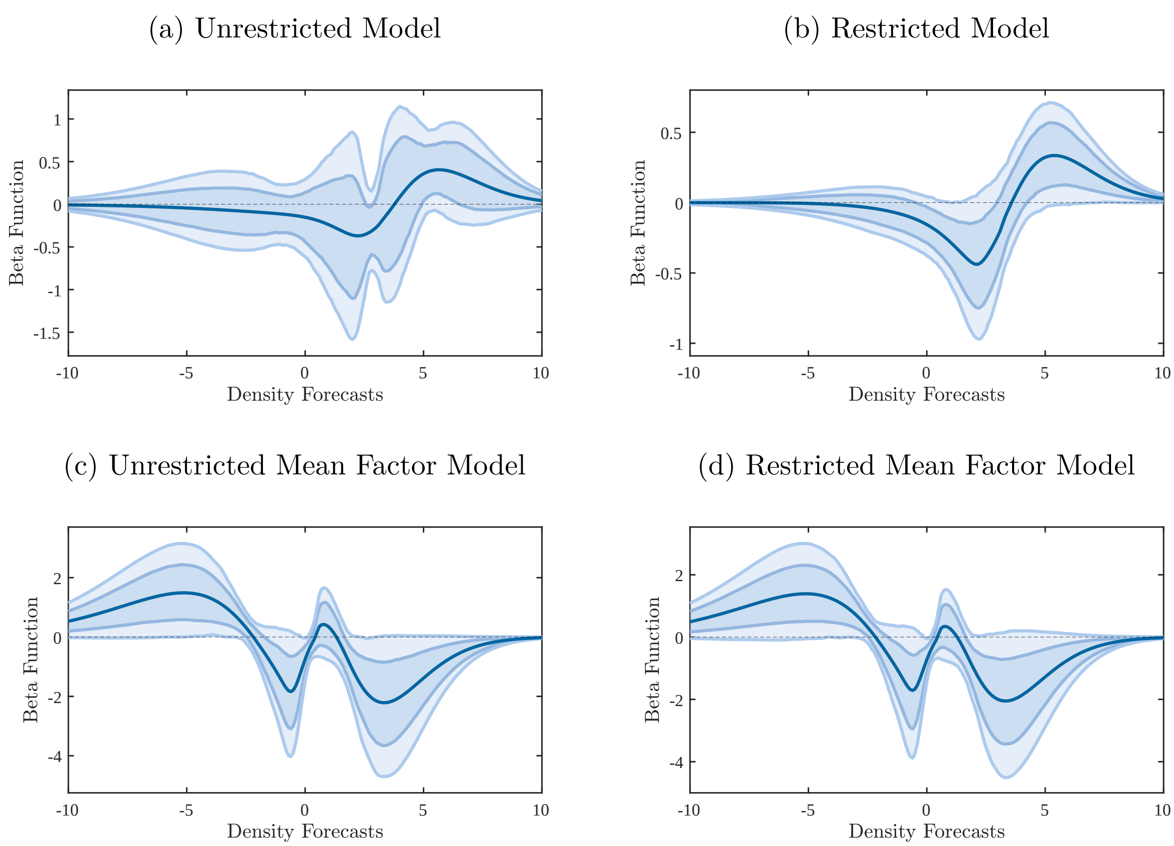
of the domain of each function being statistically significant. The more complex shape is due to more factors included as covariates in the functional regression. Note that these two functional coefficients have similar shapes with the main difference a shift to the left of the x-axis for the specification $r\mu r$ relative to specification rr. This shift is explained by the mean factor separated from the resulting centered density in Panel (d). Overall, the statistically significant region in both functions states that point forecast between 2.5% and 4% should be adjusted downwards. This recommendation holds under the restriction of full weight to the point forecast, $m_{t+h|t}$, i.e. under the assumption it is correct on average.

For one-quarter-ahead forecasts, the estimates reported in Table A2 for $\hat{\alpha}$ suggest point forecast, $m_{t+h|t}$, weights of 0.94 and 0.63 in our unrestricted specifications, ur and $u\mu r$, respectively. Relative to one-year-ahead estimates, the estimates of the weights for the point have increased substantially and the null hypothesis of full weight cannot be rejected given the confidence intervals.

The mean factor, $\left(\mu\left(f_{t+h|t}\right)\right)$, has a positive and statistically significant coefficient of 3.08 in specification $u\mu r$ and 2.67 in specification $r\mu r$. Like before, the magnitude of these coefficients should be interpreted as a weight. Additionally, the simple correlation coefficient between the point forecast, $m_{t+h|t}$, and the mean factor, $\mu(f_{+th|t})$, is 0.64, larger than from the one-year-ahead estimates.

In contrast to one-year-ahead forecasts, for one-quarter-ahead functional regressions, the

Figure A2: Beta Function Estimates For One-Quarter-Ahead GDP Growth Forecasts



Note: The figure shows the estimated functional coefficients in the respective functional regression with sample 1973Q1-2024Q3. The density forecasts are computed using linear quantile regressions of the one-quarter-ahead real GDP growth on current real GDP growth and NFCI; then, the resulting quantiles are used to fit a skewed t-distribution. We report 68% and 90% confidence bands computed using 1,000 bootstrapped samples following the algorithm detailed in Appendix A.

unrestricted specifications required more factors. Forecasting a shorter horizons usually requires a more flexible model. It is captured here by including more factors, as chosen by the BIC.

However, most of the factors are not statistically significant as shown in Table A2. Only the first leading factor is statistically significant, but barely at the 90% confidence level for the $u\mu r$ specification. This is reflected in the lack of statistical significance in all the domain of the estimated functional coefficients shown in Figure A2.

It seems that the point forecast has more relative predictive content for one-quarter-ahead than for one-year-ahead forecasts as represented by the higher weight estimates. Moreover, the predictive content of the density forecast is concentrated in the mean factor for the one-quarter-ahead forecast. Since the mean factor is highly correlated with the point forecast, there is overlapping in the predictive content that could lead to increases in the forecasting error variance for the one-quarter-ahead horizon. In contrast, for the one-year-ahead horizon, the density forecast provides information useful for forecasting that is related to recession and expansion periods. It does not seem to be entirely contained in the mean factor.

Appendix D Real-Time Forecasting Results

Table A3 shows the results from our OOS forecasting evaluation for our functional regression specifications ur, rr, $u\mu r$, and $r\mu r$ with vintage data for forecasting horizons $h = 1$ and $h = 4$. We compute the bias, variance, and the root mean squared (forecasting) error (RMSE) relative to those of the point forecast, ($m_{t+h|t}$), in the denominator. It means that improvements are measured with values below one. The estimation and forecasting procedure follows a rolling window approach as detailed in Section 4.

We also report results from two simple approaches using regression and OLS estimation. These regressions could be regarded as a direct forecasting procedure since we regress the target variable, y_{t+h}, on specific covariates. We want to address whether our method using density forecast improves relative to simple regression alternatives. First, we use the point forecast together with the covariates used for density forecasting, i.e. values of real GDP growth rate and the NFCI, but here we use the covariates directly when estimating a simple regression for the target variable. Second, we compute the inter-quartile range, i.e. the 75-percentile value minus the 25-percentile value, from the density forecasts, and use it as a covariate together with the point forecast to compute a direct forecast with simple regression for the target variable.

Table A3 shows that the improvement of point forecasts using density forecasts is possible, but not in general for our application. Panel (a) shows that our functional regression specifications without separating the mean factor have larger relative bias and variance for one-quarter-ahead GDP growth forecasts (see specifications ur and rr). When we separate the mean factor, the relative bias is smaller in both the $u\mu r$ and the $r\mu r$ specifications, but the relative variance is larger. The resulting mix was not enough to improve the point forecast, $m_{t+h|t}$, as it is evident in the RMSE reported being larger than 1. The Diebold-Mariano (DM) test null hypothesis of the same forecasting performance was rejected for specifications ur and rr, but not for $u\mu r$ and $r\mu r$ specifications. It means that the ur and rr specifications are worst than the point forecast, whereas the $u\mu r$ and $r\mu r$ specifications are relatively similar. Using directly the covariates is similar to the point forecast and including the inter-quartile range is worst. These results reflect the findings we discussed previously that the predictive content of the density forecast is concentrated in the mean factor, which is highly correlated with the point forecast, for the one-quarter-ahead forecasting horizon.

Panel (b) of Table A3 shows the respective measures for one-year-ahead GDP growth forecasts. Here, all our proposed functional regression specifications have less bias, being ur and $u\mu r$ the best with only 26% and 21%, respectively, the bias of $m_{t+h|t}$. Nonetheless, the variance is smaller by 3% only for the ur specification. These result in a decrease of

Table A3: Forecast Performance Comparison

(a) One-Quarter-Ahead

	Relative to SPF (in denominator)				
	Bias	Var	RMSE	*DM* test	K
ur	1.1693	1.3313	1.1541	2.0423**	2
rr	1.1717	1.1407	1.0696	1.8408*	1
$u\mu r$	0.3716	1.4506	1.1964	1.2539	3
$r\mu r$	0.5457	1.1100	1.0479	1.2731	3
Covariates	0.6432	2.3912	1.5370	1.0322	0
Interquantile	1.4150	1.0834	1.0473	1.6843*	0

(b) One-Year-Ahead

	Relative to SPF (in denominator)				
	Bias	Var	RMSE	*DM* test	K
ur	0.2551	0.9657	0.6640	-3.8618***	2
rr	0.2616	1.0456	0.6901	-3.3382***	1
$u\mu r$	0.2108	1.6137	0.8362	-1.0245	1
$r\mu r$	0.3029	1.0411	0.6985	-3.5177***	2
Covariates	0.2702	1.5115	0.8205	-1.2138	0
Interquantile	0.3359	1.0329	0.7048	-3.7011***	0

Note: The table shows out-of-sample forecast performance measures of the forecast (fitted) values from the functional regressions relative to the point forecast (SPF), $m_{t+h|t}$, for the target variable and forecasting horizon h. The out-of-sample is 2004Q1-2024Q3. The functional regressions estimation sample is 1973Q1-2003Q4 for GDP growth in the first rolling window. ur means unrestricted functional regression specification, rr means restricted functional regression specification, $u\mu r$ means unrestricted with mean factor functional regression specification, and $r\mu r$ means restricted with mean factor functional regression specification. The bias, variance, and root mean squared (forecasting) error (RMSE) are reported in proportion to those of the point forecast (SPF), $m_{t+h|t}$, which is included in the denominator. *DM* test is the Diebold-Mariano test statistic with the point forecast (SPF), $m_{t+h|t}$, as benchmark. ***, **, and * mean the respective forecast (fitted) value from the functional regression specification rejects the null hypothesis of equivalent performance at the 1%, 5%, and 10% significance level, respectively. A negative (positive) *DM* number means smaller (larger) errors relative to $m_{t+h|t}$. K is the number of functional principal components used in the respective functional regression as an approximation of the density forecasts. K was chosen with the BIC using the first sample in the rolling window.

the RMSE of 33% relative to $m_{t+h|t}$ for our ur specification, which is the best in terms of accuracy. The associated *DM* test statistic being negative and with significance levels below 1% show that the improvement is statistically significant for this case. Using the covariates directly does not provide improvement relative to the point forecast. In contrast, using the inter-quartile range from the density forecast together with the point forecast provides improvement of 29% in the RMSE due to a decrease in the bias. This confirms that the predictive content of the density forecast for one-year-ahead forecasting horizon comes from the information beyond the mean factor.

Our results suggest that the best way to improve the SPF's (consensus) average point forecast for one-year-ahead GDP growth is by using it jointly with density forecast. There is no improvement for one-quarter-ahead forecasts.

It seems that the improvement comes from exploiting the predictive content from the density forecast related to recession and expansion periods for one-year-ahead forecasts (see Figure 2). The functional regression coefficient estimate from the *ur* specification, shown in Figure A1, suggests that the SPF's consensus (average) forecast tends to be conservative, i.e., low rate forecasts should be adjusted downwards, whereas high rate forecasts should be adjusted upwards. This conservativeness is prominent for recessions and expansion periods because it is difficult to forecast those extreme events as they are far away from the expected value, i.e. the center of the distribution. Exploiting the information content from the tail of the distribution is the key aspect in the growth-at-risk literature and others similar.

Appendix E Robustness Checks

We consider one subsample case. We use estimation samples from 1973Q1 to 1994Q4 or 1973Q1 to 1994Q1 for $h = 1$ and $h = 4$, respectively, with first OOS date 1995Q1. In other words, we estimate our functional regressions with in-sample data until 1994Q4 in the first rolling window. The OOS sizes change accordingly with end date 2024Q3. It gives us a total of 123 OOS observations. It means that the OOS size is bigger for computing forecasting evaluation measures, but at the cost of including less information for obtaining coefficient estimates.

Here, we summarize the OOS forecasting evaluation results for our real-time forecasting exercise using vintage data. Table A4 shows results similar to those shown in the main document. Overall, there is less bias relative to the point forecast, $m_{t+h|t}$, for one-year-ahead GDP growth forecast in all our functional regression specifications. The variance is also smaller for our *ur* specification (see Panel (b) in Table A4) being the best one in terms of forecasting accuracy with a RMSE of around 65% relative to the point forecast. We don't see any improvement for one-quarter-ahead forecasts.

Another useful robustness check is to assess whether our coefficient estimates from our functional regressions have similar magnitude and sign when we use the different ways of obtaining the density forecast. We show those results in the Supplementary Material. We focus only in one-year-ahead for GDP growth forecasting and our *ur* specification because our results suggest statistically significant improvements for them.

The coefficient value for the estimated weight of of the point forecast, $m_{t+h|t}$, is robust being around 0.39 with the different ways to obtain the density forecasts considered in the paper with revised data. Also, the associated confidence intervals are similar. Moreover, the shape of the functional coefficient estimate, $\hat{\beta}$, is robust. As previously discussed, the functional coefficient suppresses smaller forecasts less than about 2.5%, while increases the occurrence of larger forecasts above 2.5%. This predictive content of the density forecast is associated with recession and expansion periods in all these cases.

Table A4: Forecast Performance Comparison With In-Sample Until 1994Q4
Real-Time Forecasting With Vintage Data

(a) One-Quarter-Ahead

	Relative to SPF (in denominator)				
	Bias	Var	RMSE	*DM* test	K
ur	0.8978	1.1446	1.0692	1.4644	3
rr	1.1372	0.9742	0.9877	-0.3565	2
$u\mu r$	0.2717	1.3938	1.1784	1.2175	5
$r\mu r$	0.4998	0.9813	0.9891	-0.2648	3
Covariates	0.6711	2.4183	1.5526	1.0508	0
Interquantile	2.1190	1.0678	1.0399	0.7909	0

(b) One-Year-Ahead

	Relative to SPF (in denominator)				
	Bias	Var	RMSE	*DM* test	K
ur	0.0912	0.8700	0.6463	-3.6921***	1
rr	0.1289	0.9954	0.6940	-3.0629***	1
$u\mu r$	0.0373	1.3268	0.7944	-1.4698	1
$r\mu r$	0.0496	1.5914	0.8703	-0.6531	1
Covariates	0.1429	0.8910	0.6588	-3.7409***	0
Interquantile	0.1482	1.0264	0.7065	-2.9735***	0

Note: Own elaboration with ALFRED data. Note: The table shows out-of-sample forecast performance measures of the forecast (fitted) values from the functional regressions relative to the point forecast (SPF), $m_{t+h|t}$, for the respective target variable and forecasting horizon h. The out-of-sample is 1995Q1-2024Q3. The functional regressions estimation sample is 1973Q1-1994Q4 for GDP growth in the first rolling window. *ur* means unrestricted functional regression specification, *rr* means restricted functional regression specification, $u\mu r$ means unrestricted with mean factor functional regression specification, and $r\mu r$ means restricted with mean factor functional regression specification. The bias, variance, and root mean squared (forecasting) error (RMSE) are reported in proportion to those of the point forecast (SPF), $m_{t+h|t}$, which is included in the denominator. *DM* test is the Diebold-Mariano test statistic with the point forecast (SPF), $m_{t+h|t}$, as benchmark. ***, **, and * mean the respective forecast (fitted) value from the functional regression specification rejects the null hypothesis of equivalent performance at the 1%, 5%, and 10% significance level, respectively. A negative (positive) *DM* number means smaller (larger) errors relative to $m_{t+h|t}$. K is the number of functional principal components used in the respective functional regression as an approximation of the density forecasts.